UNIVERSITY OF ALBERTA

Mutagenesis of Blood Group Synthesizing Glycosyltransferases

By

Ho Jun Lee

©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of

Master of Science

DEPARTMENT OF CHEMISTRY

Edmonton, Alberta

Spring, 2004

Library and Archives Canada

Published Heritage Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada

Direction du Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 0-612-96509-0
Our file Notre référence
ISBN: 0-612-96509-0

The author has granted a non-exclusive license allowing the Library and Archives Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou aturement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

ABSTRACT

Human blood group A and B glycosyltransferases are the enzymes that synthesize the blood group A and B antigens. Glycosyltransferase A (GTA) uses UDP-GalNAc as a donor and transfers the monosaccharide residue to Fucα1–2Galβ-R acceptors. Similarly, glycosyltransferase B (GTB) catalyzes the transfer of a monosaccharide residue from UDP-Gal to the same acceptors. Blood type O individuals synthesize truncated non-functional enzymes. Recently an O^2 enzyme was discovered that was a full-length form of GTA with proline 74 replaced with serine, arginine 176 replaced with glycine and glycine 268 replaced with arginine. A *cis*-AB enzyme was also discovered that was a full-length form of GTA with leucine 266 replaced with glycine and glycine 268 replaced with alanine. This enzyme exhibited A and B dual functionality. The O^2 mutant as well as the *cis*-AB mutant have been prepared by site directed mutagenesis, expressed in *E. coli*, purified, kinetically characterized and crystallized.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Professor Monica M. Palcic, for her outstanding academic support and encouragement during my graduate studies. I also thank Professor Ole Hindsgaul for providing acceptor substrates, and Professor Stephen V. Evans at the University of Victoria for his assistance with crystallography.

I thank all the group members in my research group for sharing their knowledge, especially Dr. Shuang-Jun Lin (林双君) for his helpful discussion of enzymatic synthesis, and Dr. Robert Polakowski for sharing his knowledge about enzyme kinetics. I thank Dr. Svetlana N. Borisova at the University of Victoria for growing enzyme crystals. I sincerely appreciate Zeshawn Awan at the University of Victoria for his helpful instructions about using CNS and XtalView software.

I express my deepest gratitude to my family for their constant love, encouragement, and prayers during my master's program. I also thank Lan Kim (김란), Hyun-Chan Choi (최현찬), Jonghwa Kim (김종화), Mikyung Seo (서미경), and Hyo Sun Kim (김효선) for their encouragement during my graduate studies.

TABLE OF CONTENTS

Chapter 1. Introduction **Page**

1.1 Human blood group A and B synthesizing glycosyltransferases 1
1.1.1 Topology of glycosyltransferases 3
1.1.2 General topology of GTA and GTB 7
1.1.3 Donor and acceptor substrate binding 11
1.1.4 Structure-function studies on the four critical amino acid residues 11
1.1.5 Mechanism 15

1.2 Blood group *ABO* genetics 18
1.2.1 ABO gene family 18
1.2.1.1 A alleles 18
1.2.1.2 B alleles 20
1.2.1.3 O alleles 20
1.2.1.4 *Cis*-AB alleles 23

1.3 Crystallography 26
1.3.1 Crystallization 26
1.3.2 X-ray diffraction studies 29
1.3.3 Bragg's law 29
1.3.4 Structure factor 32
1.3.5 Scattered factor 36
1.3.6 Phase problem 36
1.3.6.1. Patterson function 37
1.3.6.2. Multiple isomorphous replacement (MIR) 38
1.3.6.3. Molecular replacement (MR) 40

1.4. Overview of thesis 42

Chapter 2. Characterization of Recombinant Human Blood Group O^2 Glycosyltransferase

2.1 Introduction 48

2.2 Experimental procedures 54
2.2.1 Materials and general techniques 54
2.2.2 Cloning of P74S 56
2.2.3 Cloning of G268R 57
2.2.4 Cloning of P74S & G268R 57
2.2.5 Enzyme purification 58
2.2.6 Standard assays for GTA activity 60
2.2.6.1 P74S mutant standard assay 61
2.2.6.2 G268R mutant standard assay 62
2.2.6.3 P74S & G268R mutant standard assay 62
2.2.7 Saturating assays for GTA activity 63
2.2.7.1 P74S mutant saturating assay 63
2.2.7.2 G268R mutant saturating assay 64
2.2.7.3 P74S & G268R mutant saturating assay 64
2.2.8 Protein determination 65
2.2.9 Kinetic characterization 65
2.2.10 Crystallography 66

2.3 Results 67
2.3.1 Site-directed mutagenesis 67
2.3.1.1 P74S mutant 67
2.3.1.2 G268R mutant 69
2.3.1.3 P74S & G268R mutant 69
2.3.2 Expression and purification 71
2.3.3 Kinetic characterization 75
2.3.4 Crystal structure 80

2.4 Discussion 85

Chapter 3. Characterization of Recombinant *cis*-AB Glycosyltransferase

3.1 Introduction 96

3.2 Experimental procedures 100
3.2.1 Materials and general techniques 100
3.2.2 Cloning of L266G 103
3.2.3 Purification of L266G 105
3.2.4 Standard assay for GTA activity 107
3.2.5 Standard assay for GTB activity 108
3.2.6 Saturating assays 109
3.2.6.1 Saturating assay for GTA activity 109
3.2.6.2 Saturating assay for GTB activity 110
3.2.7 Protein determination 110
3.2.8 Kinetic characterization 111
3.2.9 Enzymatic synthesis of A and B structures by L266G 112
3.2.9.1 A-product 112
3.2.9.2 B-product 113
3.2.10 Crystallography 113

3.3 Results 114
3.3.1 Site-directed mutagenesis 114
3.3.2 Expression and purification 116
3.3.3 Kinetic characterization 119
3.3.4 Enzymatic synthesis 121
3.3.5 Crystal structure 122

3.4 Discussion 130

Chapter 4. Future work 140

LIST OF TABLES

Tables **Page**

2–1. Relative rates of transfer of the natural donor UDP-GalNAc by mutant glycosyltransferases 76
2–2. Calculated molecular weights of the mutants 78
2–3. Kinetic constants of mutant glycosyltransferases using UDP-GalNAc 79
2–4. Kinetic constants ratio for mutant glycosyltransferases using UDP-GalNAc 81
2–5. Kinetic constants of mutant glycosyltransferases using UDP-Gal 82
2–6. Kinetic constants ratio for mutant glycosyltransferases using UDP-Gal 83
2–7. Data collection and refinement results for O^2 related mutants 84
2–8. Data collection and refinement results for O^2 related mutant crystals grown in the presence of donor substrates and acceptors 86
3–1. Kinetic constants for wild-type and *cis*-AB glycosyltransferases 120
3–2. Enzymatic synthesis reactions of L266G 123
3–3. Selected ^{1}H NMR data for trisaccharide products of L266G 126
3–4. ESMS data for the trisaccharide products of L266G 127
3–5. Data collection and refinement statistics for L266G mutant grown in the unliganded and presence of donors and acceptor 128

LIST OF FIGURES

Figures **Page**

1–1. The reactions catalyzed by the human blood group A and B glycosyltransferases GTA and GTB 2

1–2. Schematic representation for a double displacement mechanism of a glycosyltransferases 4

1–3. Transmembrane topology and proteolytic processing of a glycosyltransferase 6

1–4. SETOR diagram of human blood group glycosyltransferase B (GTB) structure 8

1–5. Electron density for the glycosyltransferases in complex with H antigen and UDP 9

1–6. Coordination of the manganese ion by the DXD motif in glycosyltransferases 10

1–7. Schematic representation of significant contacts observed between GTB, acceptor and UDP 12

1–8. Model of GTA and GTB with donor and acceptor substrates 14

1–9. Schematic representations of the active sites of GTA (*A*) and GTB (*B*) 16

1–10. Schematic representation of the proposed glycosyltransferase retaining reaction mechanism 17

1–11. Schematic diagram of the genomic ABO gene 19

1–12. Common alleles at the ABO locus 21

1–13. Schematic representation of expected reading frame size for the ABO alleles with altered reading frame length 24

1–14. Schematic representation of the divergent amino acid positions among GTA, GTB and *cis*-AB enzymes 25

1–15. Schematic representation of the unit cell 28

1–16. Schematic representation of X-ray diffraction 30

1–17. Mathematical construction of crystals and Bragg's law 31

1–18. Vector diagram of structure factor 33

1–19. Schematic representation of Friedel's law 35

1–20. Contributions of vectors to the structure factor in MIR 39

1–21. Schematic diagram of the phase calculation using MIR 41

2–1. The reactions catalyzed by the human blood group A and B glycosyltransferases GTA and GTB 49
2–2. Schematic diagram of the divergent amino acid positions for GTA and GTB 50
2–3. Schematic representation of significant contacts observed between GTB, acceptor, and UDP 52
2–4. Schematic diagram of the amino acid positions among GTA, GTB and O^2 glycosyltransferase 53
2–5. Agarose Gel of PCR reaction for the P74S mutation 68
2–6. Agarose Gel of digested mutants 68
2–7. Agarose Gel of the first PCR fragments for the G268R product 70
2–8. Agarose Gel of the second PCR reaction for the G268R product 70
2–9. Agarose Gel of the first PCR reaction of P74S & G268R product 72
2–10. Agarose Gel of digested P74S & G268R product 72
2–11. SDS-PAGE result for the P74S mutant 73
2–12. SDS-PAGE result for the G268R mutant 73
2–13. SDS-PAGE result for the P74S & G268R mutant (double mutation) 74
2–14. The electron density map of the active site of the O^2 enzyme 87
2–15. The structure of the active site of the O^2 enzyme with donor and acceptor substrates 91
3–1. The reactions catalyzed by the human blood group A and B glycosyltransferases 97
3–2. Schematic representation of the divergent amino acid positions for GTA and GTB 98
3–3. Schematic diagram of the changes in amino acids that give *cis*-AB enzymes 101
3–4. Agarose Gel of the first PCR fragments for the L266G mutant 115
3–5. Agarose Gel of the second PCR reaction for the L266G mutant 115
3–6. Agarose Gel of digested L266G mini plasmid 117
3–7. SDS-PAGE result for L266G mutant 118
3–8. ^1H NMR spectra (in D_2O) of A-product

(GalNAcα1–3[Fucα1–2]Galβ-O-$(CH_2)_7CH_3$) 124
3–9. ^{1}H NMR spectra (in D_2O) of B-product
(Galα1–3[Fucα1–2]Galβ-O-$(CH_2)_7CH_3$) 125
3–10. The electron density map of active cleft of the L266G mutant 129
3–11. The electron density map of the active cleft of the L266G enzyme (AAGlyB) with donor substrates 133
3–12. Schematic representation of the active site of the *cis*-AB enzyme (AAGlyB) with donor substrates 134
4–1. The nucleotide sequence of the ABO blood group alleles which has not yet been characterized 142
4–2. The amino acid sequence of the ABO blood group alleles which has not yet been characterized 143

LIST OF ABBREVIATIONS

aa	amino acid
bp	base pair
BSA	bovine serum albumin
Ci	Curie
CNS	crystallography and NMR software
C-terminal	carboxy terminal
Da	Dalton
cDNA	complementary deoxyribonucleic acid
dpm	disintegrations per minute
DTT	dithiothreitol
E. Coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
ER	endoplasmic reticulum
ESMS	electrospray mass spectrometry
Fuc	fucose
Gal	galactose

GalNAc	*N*-acetylgalactosamine
GT	glycosyltransferase
GTA	α1-3 *N*-acetylgalactosaminyltransferase
GTB	α1-3 galactosyltransferase
Hz	Hertz
IgG	immunoglobulin G
IPTG	isopropyl-1-thio-β-D-galactopyranoside
k_{cat}	catalytic constant
kDa	kilodalton
Km	Michaelis-Menten constant
LB	Luria-Bertani broth
LFT	load-flow-through
MHz	megahertz
MIR	multiple isomorphous replacement
MOPS	3-(*N*-morpholino) propanesulfonic acid
MR	molecular replacement
NMR	nuclear magnetic resonance

nt	nucleotide
N-terminal	amino terminal
OD	optical density
PCR	polymerase chain reaction
ppm	parts per million
rpm	revolutions per minute
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
TB	terrific broth
TLC	thin layer chromatography
UDP	uridine 5'-diphosphate
UDP-Gal	uridine 5'-diphosphogalactose
UDP-GalNAc	uridine 5'-diphospho-*N*-acetylgalactosamine
V_{max}	maximal reaction velocity
WFT	wash-flow-through

Chapter 1. Introduction

1.1 Human blood group A and B synthesizing glycosyltransferases

The ABO(H) histo-blood groups are clinically important in blood transfusion and organ transplantation. They were first discovered by Landsteiner in 1900 (1, 2) and later shown to be carbohydrates (3, 4) synthesized by glycosyltransferases (5, 6). Extensive genetic, structural and kinetic studies have been carried out on both the blood group A and B synthesizing glycosyltransferases.

The human A and B blood group glycosyltransferases are enzymes that catalyze the transfer of a monosaccharide residue from the nucleotide donors, UDP-GalNAc or UDP-Gal, to Fucα1–2Galβ-R (H)-terminating acceptors (3, 4). The blood group A synthesizing α1–3 *N*-acetylgalactosaminyltransferase (GTA, EC 2.4.1.40) transfers GalNAc from UDP-GalNAc to acceptor, producing the A antigen GalNAcα1–3[Fucα1–2]Galβ-R. The blood group B synthesizing α1–3 galactosyltransferase (GTB, EC 2.4.1.37) transfers Gal from UDP-Gal to acceptor, producing the B antigen Galα1–3[Fucα1–2]Galβ-R (Fig. 1–1). These reactions occur with retention of anomeric center

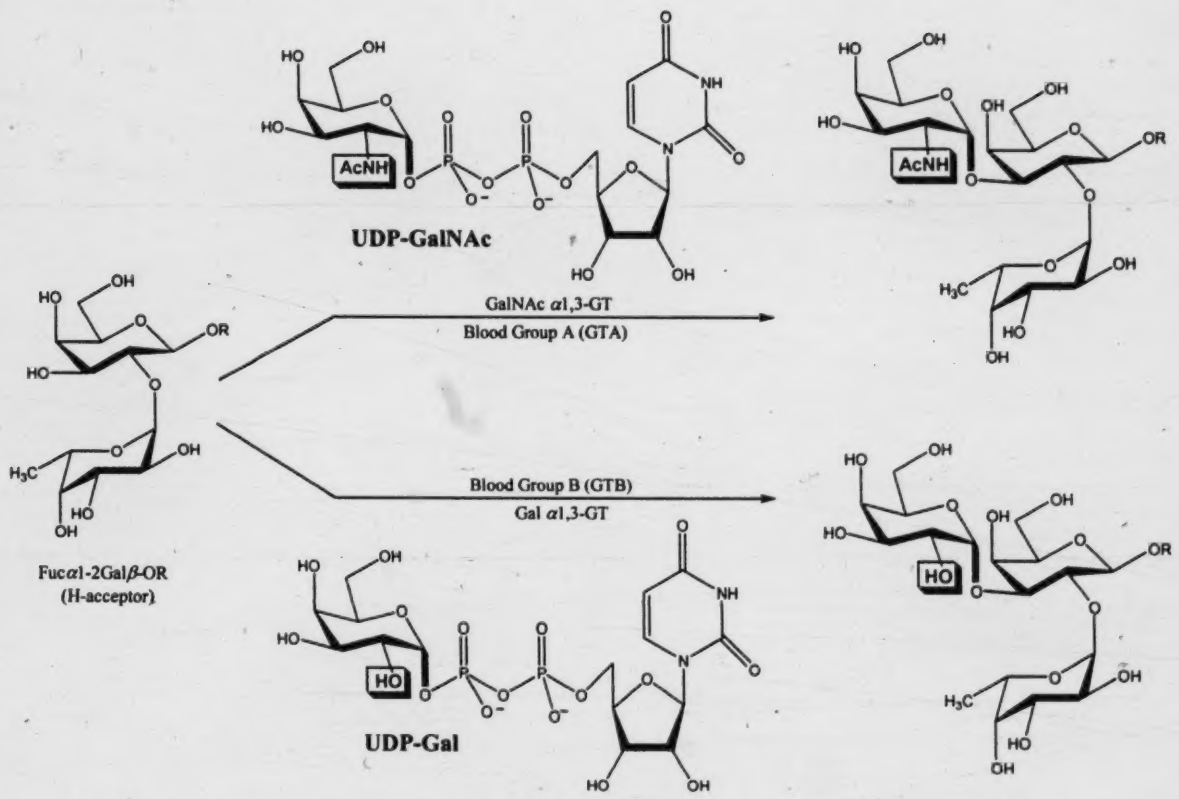

FIG. 1–1. **The reactions catalyzed by the human blood group A and B glycosyltransferases GTA and GTB.**

of the transferred monosaccharide, consistent with a double displacement mechanism where a glycosyl-enzyme intermediate forms prior to reaction with the acceptor (5) (Fig. 1–2).

GTA and GTB are highly homologous enzymes differing at only four critical amino acids out of a total of 354 residues (6). Alteration of these four crucial amino acid residues, Arg-176 → Gly, Gly-235 → Ser, Leu-266 → Met, and Gly-268 → Ala, converts the specificity from GTA to GTB. The roles of each of these amino acids in determining the reaction specificity have been elucidated and will be discussed in section 1.1.4.

1.1.1 Topology of glycosyltransferases

The blood group synthesizing glycosyltransferases belong to the type II class of integral membrane proteins and share a common structure with other mammalian glycosyltransferases. These enzymes are located in the membrane of the endoplasmic reticulum (ER) and the Golgi apparatus. They have a short N-terminal cytoplasmic tail, a transmembrane signal-anchor domain, a stem region and a C-terminal catalytic domain

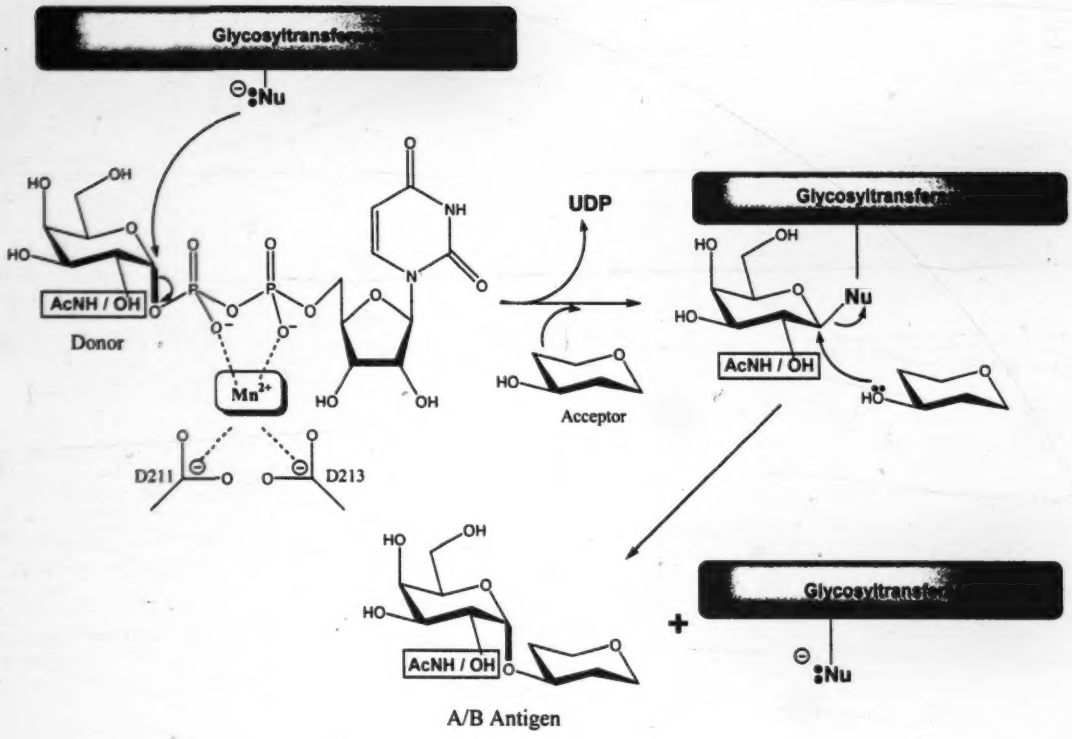

FIG. 1–2. **Schematic representation for a double displacement mechanism of a glycosyltransferase.**

(Fig. 1–3) (7). The transmembrane signal-anchor domain (8) acts as both an uncleavable signal peptide and as a membrane-spanning region and orients the catalytic domain of the glycosyltransferase into the lumen of the ER and Golgi (Fig. 1–3). The enzymes in the ER lumen are transported to the Golgi complex, which receives and exports proteins to various destinations via transport vesicles. The stem region shown in Fig. 1–3 serves as a flexible and proteolytically sensitive linkage. After the proteolytic cleavage event, a catalytically active fragment of the glycosyltransferase is released to the extracellular milieu through the Golgi apparatus.

GTA and GTB are low abundance enzymes therefore they were cloned and expressed in *E. coli*, generating large quantities of enzyme for donor specificity analysis. Synthetic genes for GTA and GTB were constructed by Dr. N. O. L. Seto of the Institute for Biological Sciences, National Research Council of Canada in Ottawa with codon usage optimized for *E. coli*. This facilitates mutagenesis and the production of soluble truncated enzymes lacking the transmembrane domain (9, 10). GTA and GTB enzymes truncated in the stem region at amino acid 54 formed aggregates and precipitated. Removal of ten more amino acids generated soluble active enzyme that could be concentrated to 20 mg/ml for

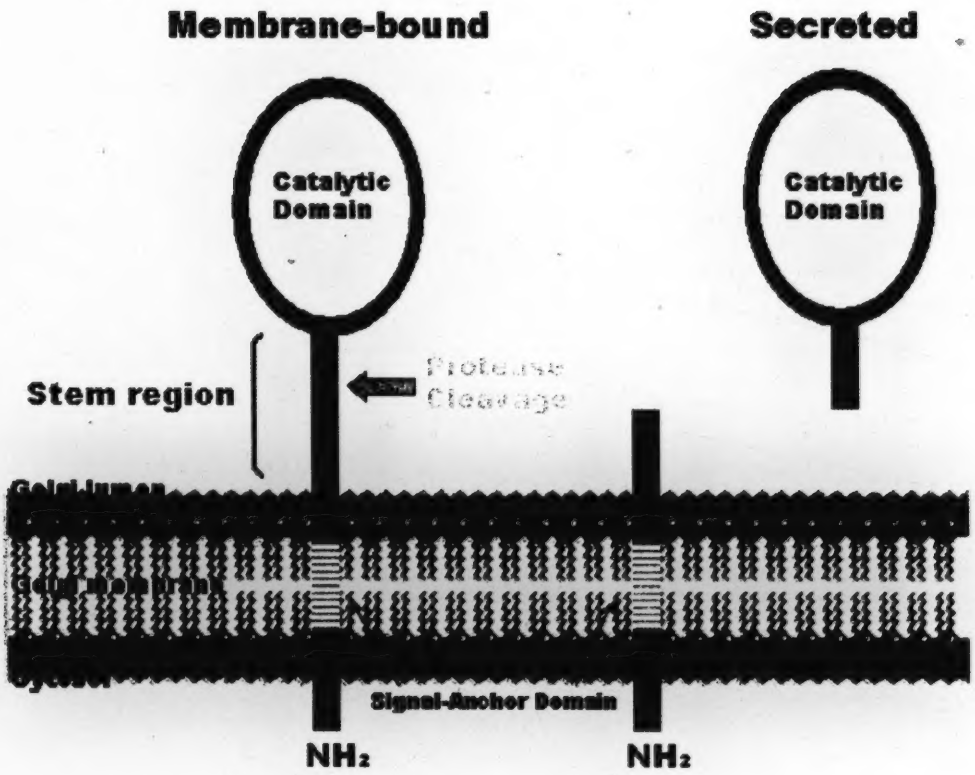

FIG. 1–3. **Transmembrane topology and proteolytic processing of a glycosyltransferase.**

X-ray crystallography studies. This is the constant that was used as a template for the production of mutant glycosyltransferases in the thesis.

1.1.2 General topology of GTA and GTB

The X-ray structures of soluble GTA and GTB (amino acid residues 63–354) show they are organized in two domains separated by an active site cleft (11). The cleft is approximately 13 Å wide and contains the four critical amino acid residues that are different in GTA and GTB (Fig. 1–4). Unambiguous electron density is observed for the H-antigen disaccharide substrate and UDP complexed to GTB (Fig. 1–5). The structure shows that the N-terminal domain (residues 96–227, 338–345, β3–β7, β8, β13 and α1–α3) recognizes the nucleotide donor. The disaccharide acceptor-binding site is formed by residues in the C-terminal domain (residues 228–237, β7, β10, β11, β2, α4 and α5) in combination with UDP. A DXD motif (12) comprised of residues Asp-211, Val-212 and Asp-213 is located in the middle of cleft, where it coordinates to a Mn^{2+} ion that has been suggested to have a role in catalysis (13). Both Asp-211 and 213 interact with the Mn^{2+} with bidentate coordination (Fig. 1–6).

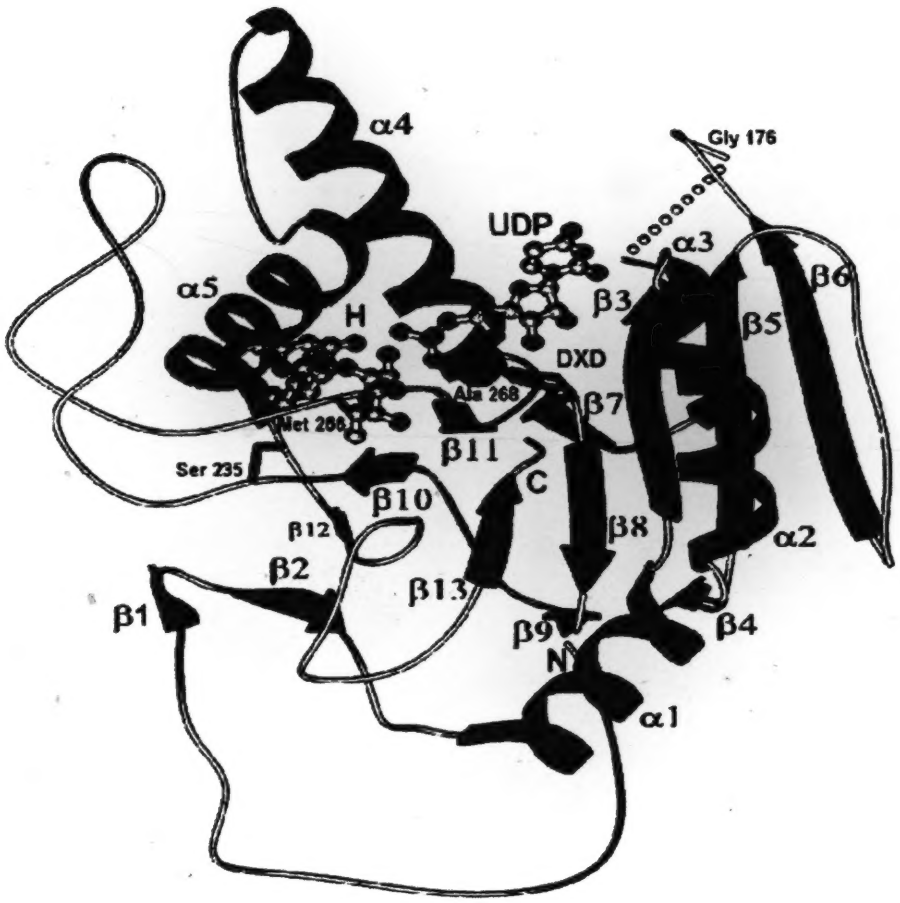

FIG. 1–4. **SETOR (15) diagram of human blood group glycosyltransferase B (GTB) structure.** The enzyme consists of two domains separated by a wide central cleft that contains the active site an UDP and the H-antigen acceptor (ball-and-stick model), as well as the DXD motif (Asp-211 and Asp-213), which coordinates to a Mn^{2+} ion (*magneta*). The N-terminal region is represented by *dark blue* and *magneta* and the C-terminal region is represented by *light blue* and *green*. GTA and GTB differ in only four critical amino acid residue positions, Arg/Gly-176, Gly/Ser-235, Leu/Met-266, Gly/Ala-268 shown in brown (From Patenaude, S. I., Seto, N. O. L., Borisova, S. N., Szpacenko, A., Marcus, S. L., Palcic, M. M., and Evans, S. V. (2002) *Nat. Struct. Biol.* **9**, 685–690).

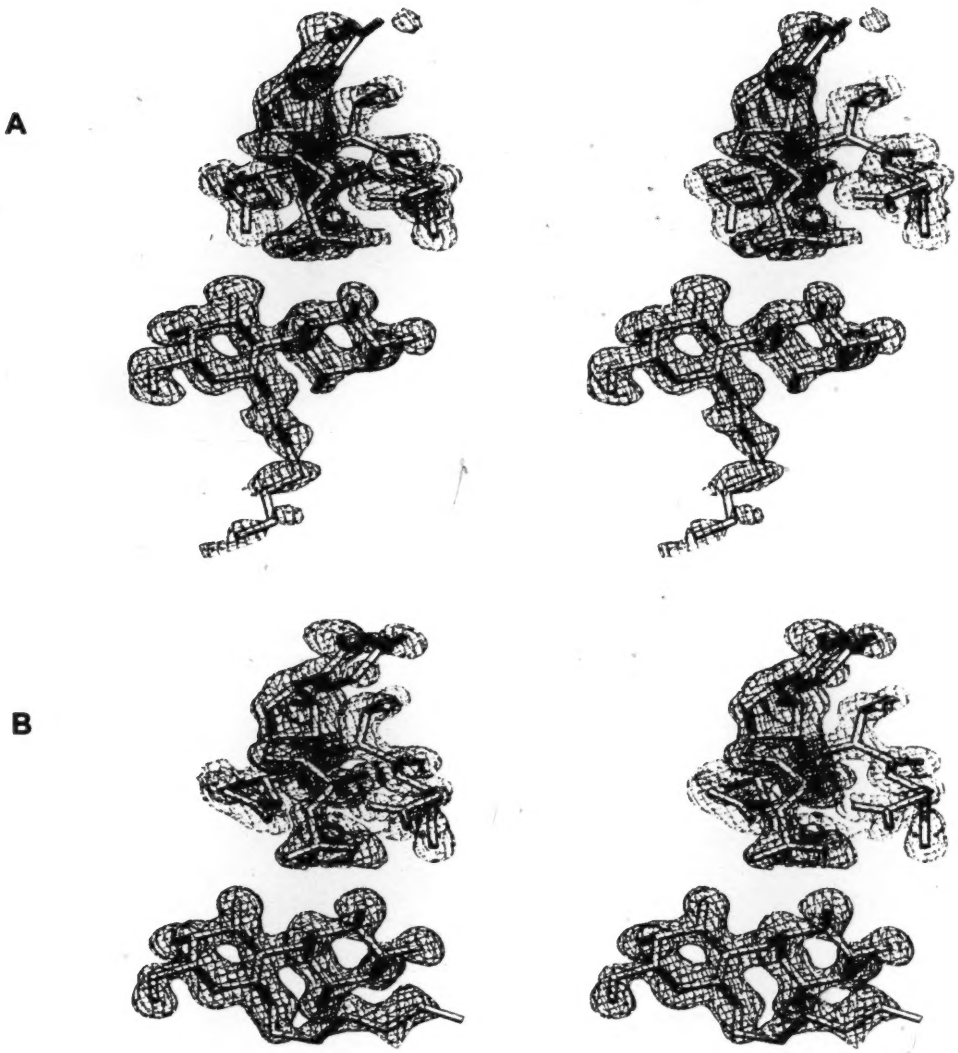

FIG. 1–5. **Electron density for the glycosyltransferases in complex with H antigen and UDP.** Electron density for GTA (*A*), GTB (*B*), the DXD motif (12), (Asp-211, Val-212, Asp-213) and manganese is shown in *red*, UDP in green and H antigen disaccharide acceptor in *blue* (From Patenaude, S. I., Seto, N. O. L., Borisova, S. N., Szpacenko, A., Marcus, S. L., Palcic, M. M., and Evans, S. V. (2002) *Nat. Struct. Biol.* **9,** 685–690).

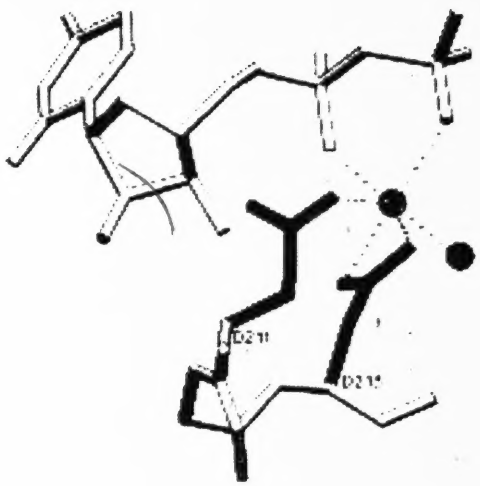

FIG. 1–6. **Coordination of the manganese ion by the DXD motif in glycosyltransferases.** The GTB DXD coordination of manganese (*red*) is typical of retaining glycosyltransferase. Both Asp-211 and 213 (*green*) contact the manganese ion with bidentate coordination. The main chain and UDP are represented in *white*, DXD residues in *green* and water in *blue* (From Patenaude, S. I., Seto, N. O. L., Borisova, S. N., Szpacenko, A., Marcus, S. L., Palcic, M. M., and Evans, S. V. (2002) *Nat. Struct. Biol.* **9,** 685–690).

1.1.3 Donor and acceptor substrate binding

The protein-substrate contacts from the X-ray structure of GTB are shown in Fig. 1–5 and 1–7. The H disaccharide acceptor interacts with the enzyme through hydrogen bonds with the side chains of His-233 and Glu-303 and the galactose O-4 hydroxyl group. The Gal O-6 hydroxyl group hydrogen bonds to Thr-245 and the Fuc O-4 hydroxyl to Asp-326. In GTB the acceptor makes a van der Waals contact with one of the critical residues Met-266, which is involved in the selection of donor. Both sugar residues of the disaccharide acceptor make strong hydrogen bonds with the β-phosphate group of the UDP, indicating that the binding of the donor is a crucial preliminary step in the formation of the acceptor-binding site. Tyr-126 interacts with the α-phosphate group of the UDP and Phe-121 can hydrogen bond with the hydroxyl group of the ribose in UDP. The contacts are identical for GTA, except there is no interaction between Leu-266 and acceptor (11).

1.1.4 Structure-function studies on the four critical amino acid residues

The first of the critical amino acids (Arg/Gly-176) is located the furthest away from

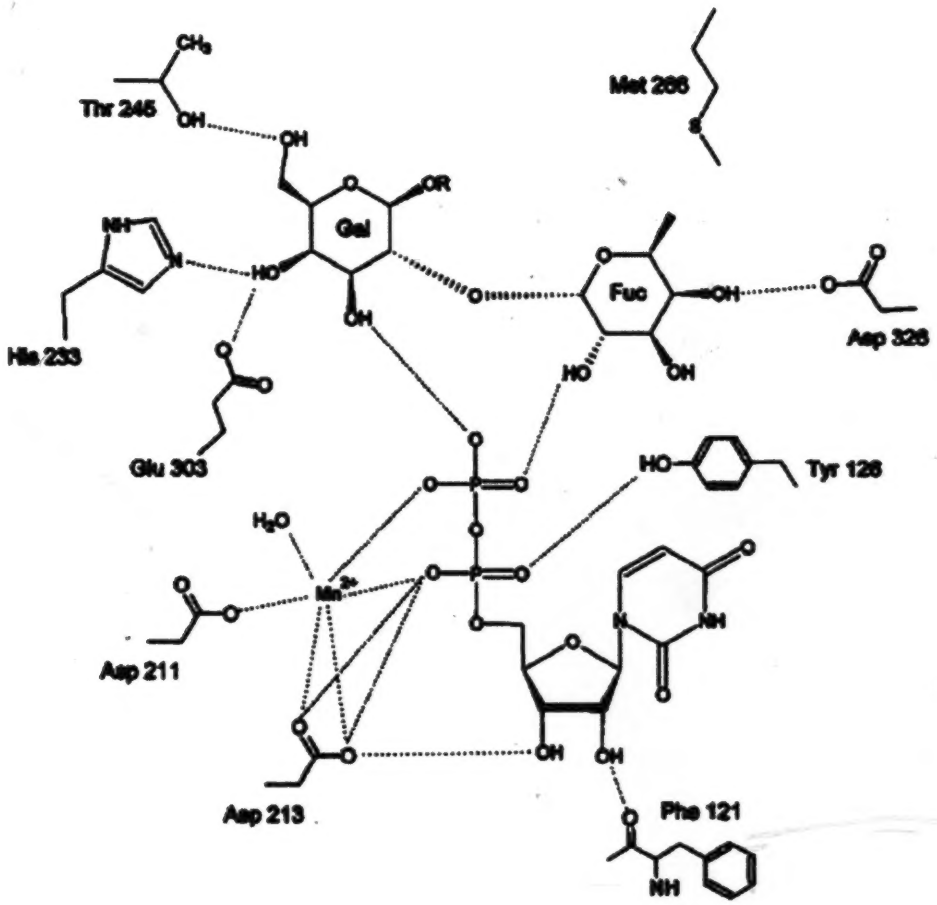

FIG. 1–7. **Schematic representation of significant contacts observed between GTB, acceptor and UDP** (From Patenaude, S. I., Seto, N. O. L., Borisova, S. N., Szpacenko, A., Marcus, S. L., Palcic, M. M., and Evans, S. V. (2002) *Nat. Struct. Biol.* **9,** 685–690).

the substrate-binding site. Even though these residues are unlikely to contact the donor substrate, mutation from Arg to Gly in wild-type GTA caused an 11-fold increase in enzyme turnover (k_{cat}) without affecting the donor specificity. This observation suggests a role for these residues in turnover or product release. The second critical amino acids (Gly/Ser-235) are not positioned to contact the donor directly, although they force the aliphatic tail on the acceptor to assume different conformations in GTA and GTB (Fig. 1–8). The different conformations might explain the 3-fold decrease of acceptor K_m for GTA compared to GTB (10, 14). It has been suggested that Gly/Ser-235 might select between different structures displaying the H antigen acceptor (11). Only two (Leu/Met-266 and Gly/Ala-268) of the four critical residues occupy positions in the active sites of GTA and GTB. Since both enzymes hydrolyse their UDP-saccharide donors, even in the absence of acceptor, structures for complete donor complexed to enzyme could not be obtained. However, the excellent electron density observed for UDP donors (Fig. 1–5) allows for modeling of the monosaccharide of each donor (Fig. 1–9). Therefore, they control reaction specificity. Leu/Met-266 is positioned to accommodate either the acetamido or hydroxyl group thereby distinguishing between the A donor (UDP-GalNAc) and the B donor

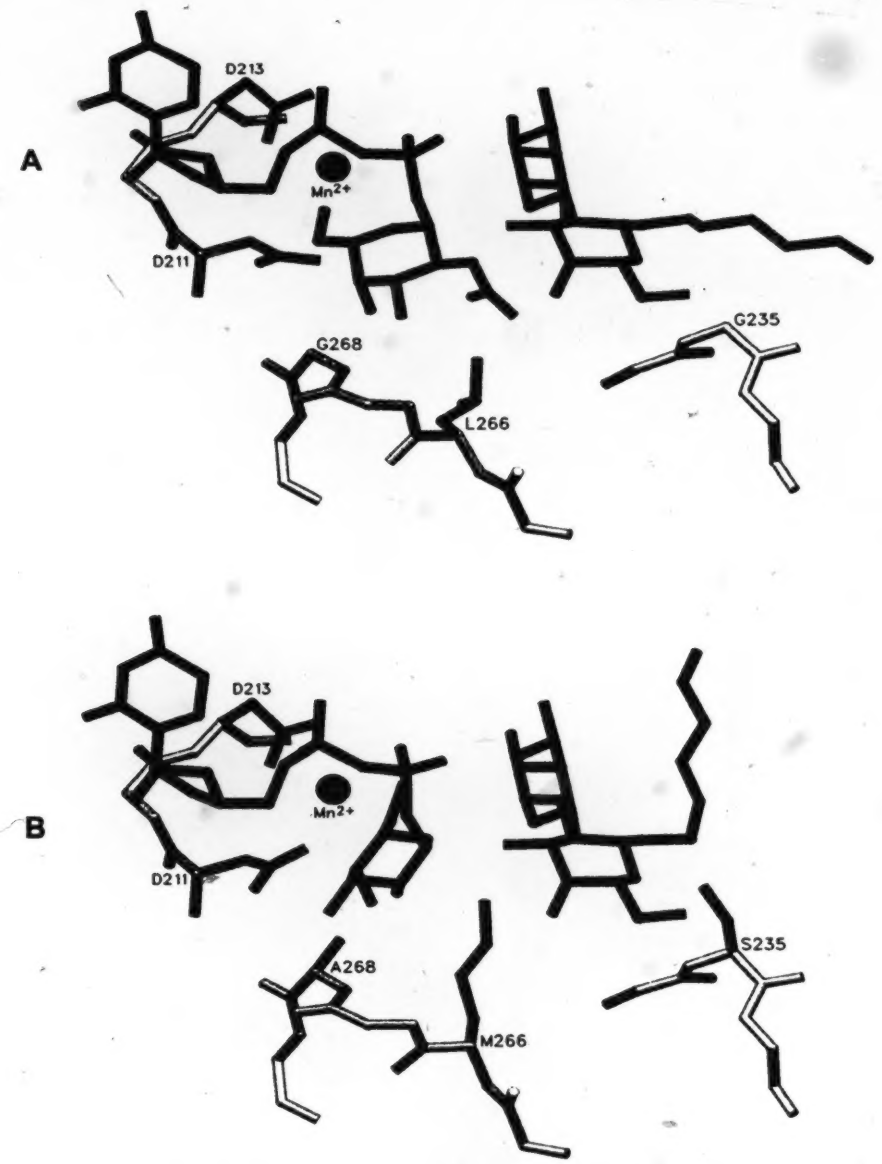

FIG. 1–8. **Model of GTA and GTB with donor and acceptor substrates.** Location of the substrates (*green*) in the GTA (*A*) and GTB (*B*) binding sites with respect to 3 of the 4 critical residues (*blue*) that discriminate enzymatic activities. Only residues at amino acid 266 and 268 are positioned to interact with the donor substrate. The GalNAc and Gal sugar moieties were modeled into the binding site by SETOR (15) (From Patenaude, S. I., Seto, N. O. L., Borisova, S. N., Szpacenko, A., Marcus, S. L., Palcic, M. M., and Evans, S. V. (2002) *Nat. Struct. Biol.* **9**, 685–690).

(UDP-Gal). This is a complementary interaction where the smaller Leu-266 in GTA accommodates the larger acetamido group in UDP-GalNAc. Similarly, UDP-Gal, which has a smaller hydroxyl group, is accommodated by the larger Met-266 in GTB (14). Residues Met-266 and Ala-268 in GTB are bulkier than the corresponding residues in GTA, excluding the bulkier UDP-GalNAc. In GTA, the larger active site provided by smaller residues, makes a larger void for UDP-Gal (11) (Fig. 1–9).

1.1.5 Mechanism

A suggested mechanism for retaining blood group glycosyltransferases involves nucleophilic attack by a residue such as Glu-303 giving a glycosyl-enzyme intermediate shown as a covalent adduct in Fig. 1–10. This is followed by the transfer of monosaccharide from the enzyme to acceptor. This double displacement mechanism is consistent with retention of anomeric configuration of the transferred monosaccharide.

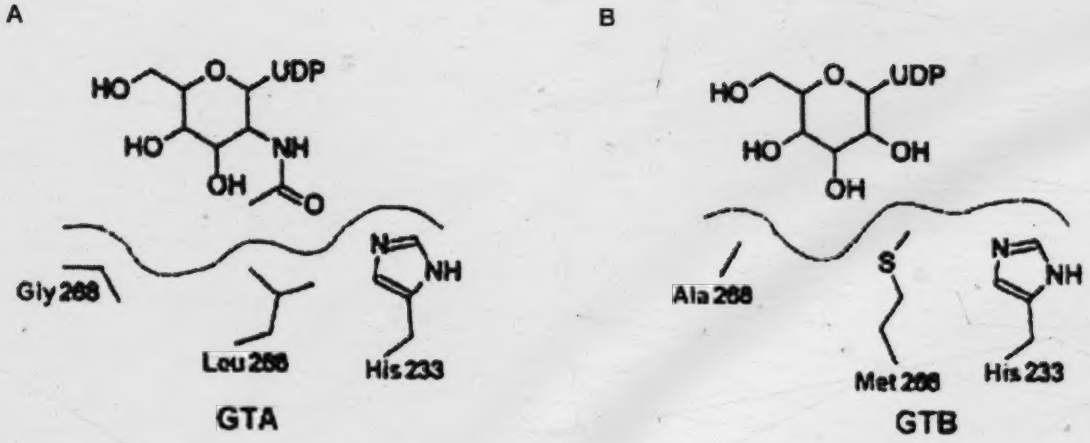

FIG. 1–9. **Schematic representations of the active sites of GTA (*A*) and GTB (*B*).** GTB can easily exclude the bulkier UDP-GalNAc since its active site cleft contains the larger residues Met-266 and Ala-268. The smaller critical residues Leu-266 and Gly-268 in GTA make a larger active site cleft that would lack complementary interaction with the smaller UDP-Gal (From Patenaude, S. I., Seto, N. O. L., Borisova, S. N., Szpacenko, A., Marcus, S. L., Palcic, M. M., and Evans, S. V. (2002) *Nat. Struct. Biol.* **9**, 685–690).

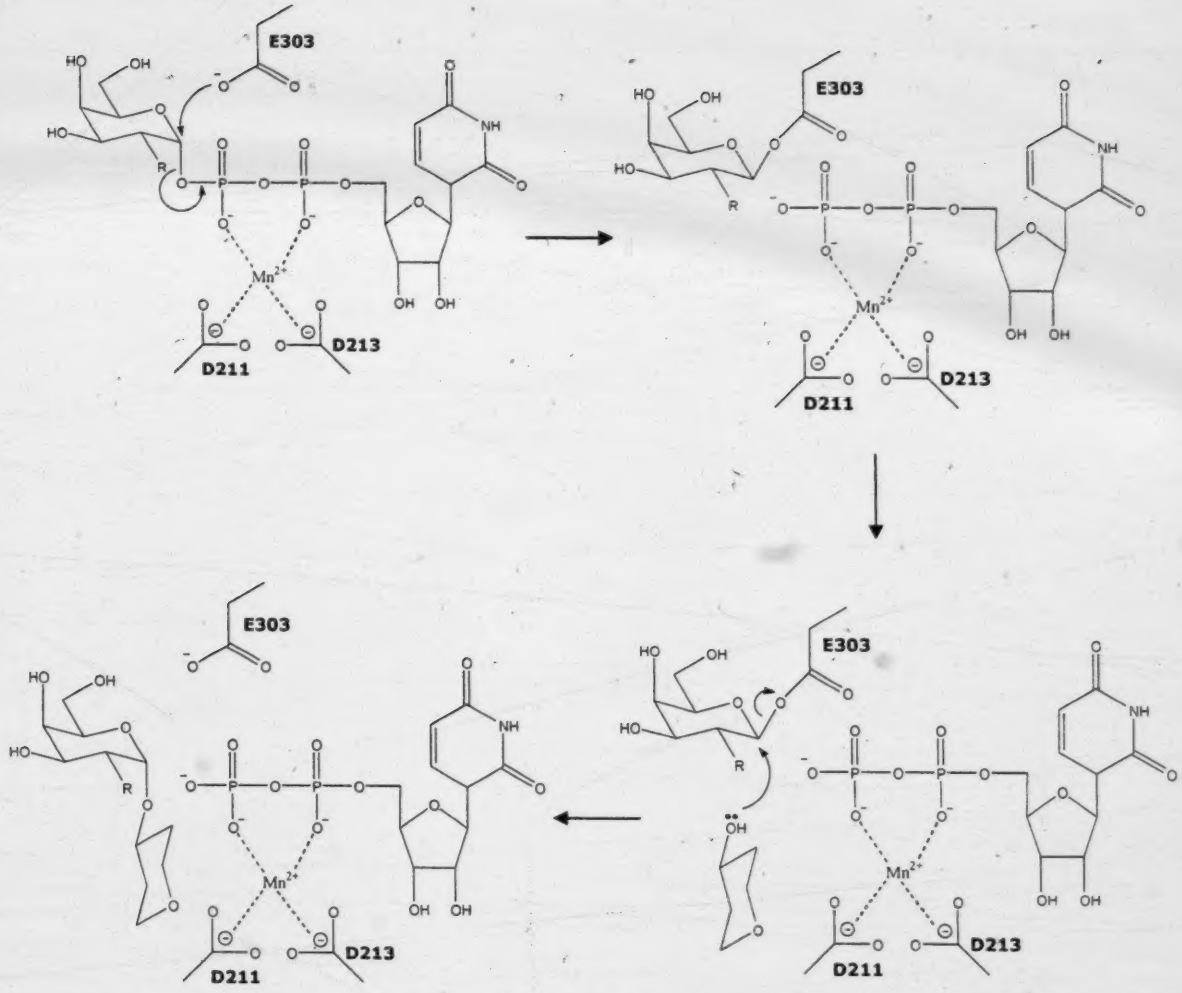

FIG. 1–10. **Schematic representation of the proposed glycosyltransferase retaining reaction mechanism.** R denotes the NHAc group in GTA and the OH group in the GTB reaction.

1.2 Blood group ABO genetics

1.2.1 ABO gene family

The ABO gene contains seven exons spanning approximately 18–20 kilobase pairs (16, 17). A schematic diagram of the gene is shown in Fig. 1–11. Studies of the ABO gene started on the translated exons, and then intron sequencing was carried out (16, 17). A submission to GenBank (AC000397) included the full sequences of exons 1 – 6 and part of exon 7, as well as all six introns. Those exon sequences indicated that their source was an A^1 allele. The cDNA of the human blood group A^1 allele contains a 1062-base pair nucleotide sequence encoding a 41-kDa enzyme (18).

1.2.1.1 A alleles

The first A allele described previously (16, 17) was the A^1 allele, which contains 1062-base pair nucleotide sequence in its cDNA. In addition to this consensus sequence there is another major A^1 allele containing a C467T mutation which causes a

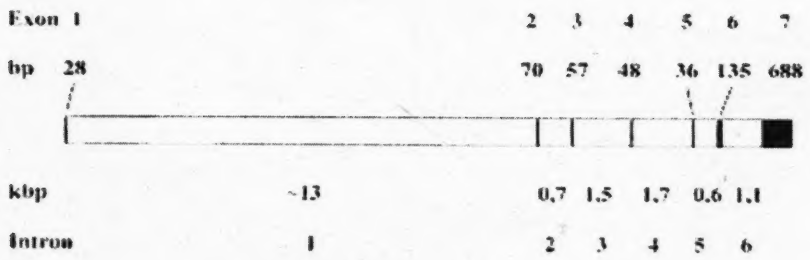

FIG. 1–11. **Schematic diagram of the genomic ABO gene.** The exons (filled rectangles) and the introns (open rectangles) are drawn to scale. The number of nucleotides in the exons and approximate number of introns are shown (From Olsson, M. L., and Chester, M. A. (2001) *Transfus. Med.* **11,** 295–313).

Pro-156 → Leu alteration. However, there is no change in enzymatic activity for this mutant (18). The A^2 allele contains the same nucleotide mutation at codon 467 and also a single nucleotide deletion in a sequence of three consecutive cytidines in nucleotide positions 1059–1061 of the A^1 allele. This deletion results in an extension of the reading frame by 64 nucleotides giving a longer A^2 protein (Fig. 1–12).

1.2.1.2 B alleles

The length of translated DNA is identical in the A and B glycosyltransferases, except for seven single nucleotide substitutions which result in four amino acid changes in the expressed protein (Fig. 1–12) (6). In addition, there is another mutation at the nucleotide at position 1096 (G1096A) beyond the stop codon that is useful for ABO genotype screening (20).

1.2.1.3 O alleles

An O allele can be any allele having a mutation that causes a loss of functionality. The blood group O^1 allele was originally defined as a consensus allele, but containing a

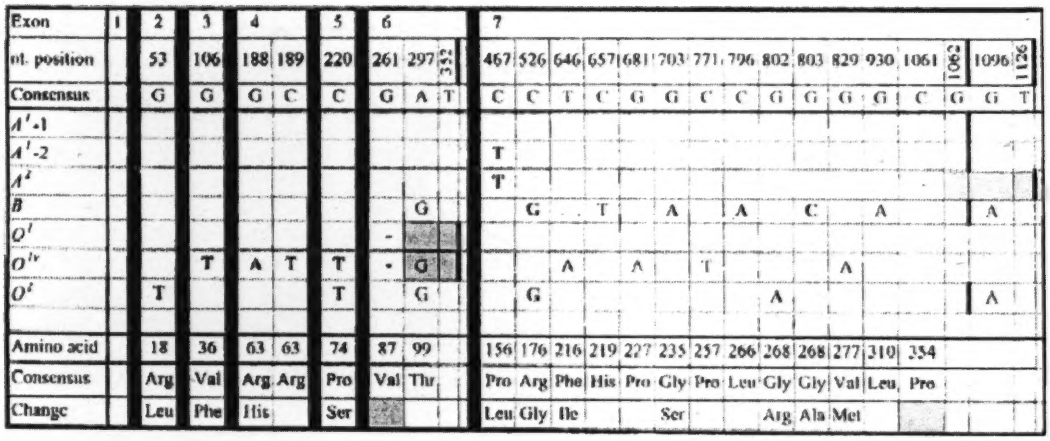
Exon 1 2 3 4 5 6 7
nt. position 53 106 188 189 220 261 297 332 467 526 646 657 681 703 771 796 802 803 829 930 1061 1063 1096 1126
Consensus G G G C C G A T C C T C G G C C G G G G C G G T
A1-1
A1-2
A2
B
O1
O1v
O2
Amino acid 18 36 63 63 74 87 99 156 176 216 219 227 235 257 266 268 268 277 310 354
Consensus Arg Val Arg Arg Pro Val Thr Pro Arg Phe His Pro Gly Pro Leu Gly Gly Val Leu Pro
Change Leu Phe His Ser Leu Gly Ile Ser Arg Ala Met

FIG. 1–12. **Common alleles at the ABO locus.** Only changes from the consensus (A^1-1) are shown. Mutations causing amino acid changes are represented by bold type. The dark vertical bars indicate the positions of the introns. The thick vertical lines indicate the end of reading frames. The shadowed area indicates a reading frame shift (From Olsson, M. L., and Chester, M. A. (2001) *Transfus. Med.* **11,** 295-313).

single nucleotide deletion at codon 261 in exon 6. This deletion causes a frame shift that changes the protein sequence after amino acid 88 and introduces a stop codon after the nucleotide at position 352. Therefore translation stops at amino acid 117 giving a non-functional truncated protein. A second O allele termed O^{1v} was observed in a COLO205 cell line (6). It contained the crucial deletion at the nucleotide 261 and a number of point mutations though exon 3–7. The O^2 allele was the first O allele that did not have the G261 deletion (21). Rather it was a new mutation at nucleotide position 802 which caused an amino acid change (Gly-268 → Arg). This alteration of the glycine to arginine presumably adversely affects nucleotide sugar binding accounting for the loss of enzyme activity (22, 23). The O^3 allele is other rare O allele without the G261 deletion. This type of O allele differs from the consensus A^1 allele by having the common A^2 deletion (C1061-) and the A^{el} allele insertion (20) of an extra guanosine in the sequence at nucleotide 798–804 (24). This double frame shift can translate a normal length of protein compared to A^1 allele, however, these changes including deletion and insertion alter the amino acid sequence from 269 to 353 (25). As a result of these alterations (amino acid 269–353) the O^3 allele shows non-functionality. O^4 is caused by a guanosine insertion at nucleotide position 87–88, resulting in a frame-shift and introducing a stop codon at codon 56. Finally, the O^5 allele

with a C322T mutation directly introduces a stop codon (25). O^4 and O^5 alleles cannot translate a full-length of gene, therefore, they show loss of activity. All of the alleles described above are shown schematically in Fig. 1–13.

1.2.1.4 *Cis*-AB alleles

Cis-AB alleles code for enzymes that have both A and B activities. The first characterization of genomic DNA from a *cis*-AB individual showed an intermediate form of the ABO gene. It corresponded to AAAB, where A and B represent the A- or B- specific nucleotides at the four positions in exon 7 of the gene that alter the amino acid composition of the expressed protein (Fig. 1–14) (26–28). Subsequently other types of *cis*-AB gene have been reported, which contain BABB, BBAB, ABAB and BAAB constellations at these four positions (27, 29). Another type of the *cis*-AB allele was discovered that was a full-length form of GTB with proline 234 replaced with alanine (30). More recently, a new type of the *cis*-AB allele was discovered that was a full-length form of GTA with leucine 266 replaced with glycine and glycine 268 replaced with alanine. This *cis*-AB allele prevails in the mouse population (31). All of the *cis*-AB enzymes described above are shown in Fig. 1–14.

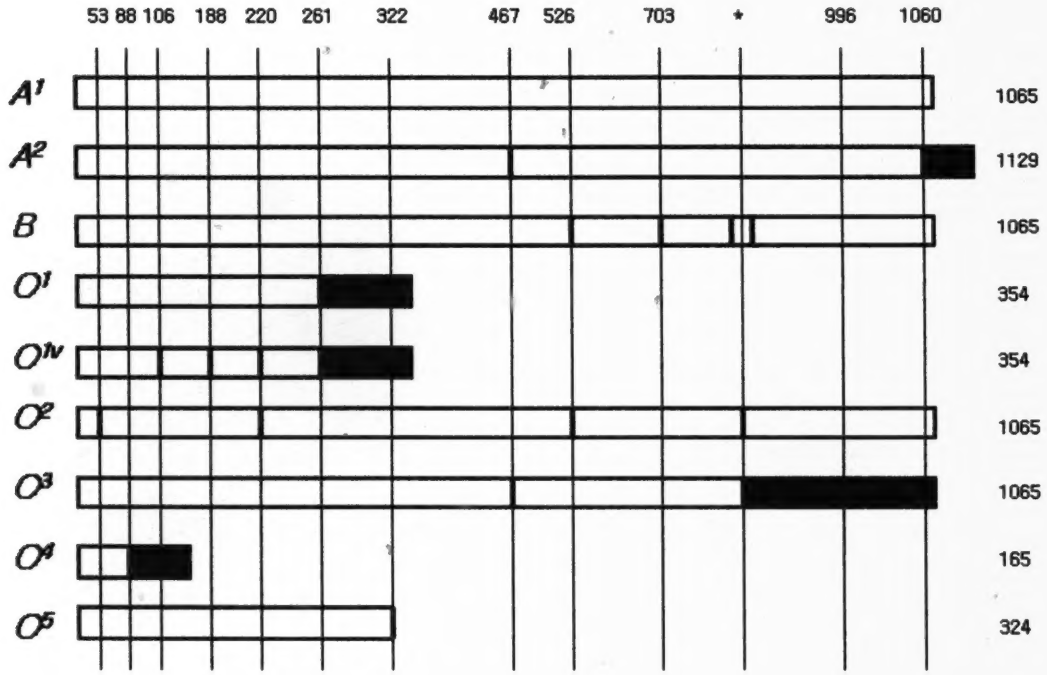

FIG. 1–13. **Schematic representation of expected reading frame size for the ABO alleles with altered reading frame length.** The open bars denote translated A^1 consensus; the filled bars indicate translated non-A^1 consensus. The reading frame size is shown at the end of bars. The solid *blue* vertical lines show nucleotide positions of mutations which lead the alterations of amino acid. Asterisk (*) represents the B allele mutation at nt 796 and 803; the O^2 allele mutation at 802; O^3 allele insertion at nt798–804 (From Olsson, M. L., and Chester, M. A. (2001) *Transfus. Med.* **11,** 295–313).

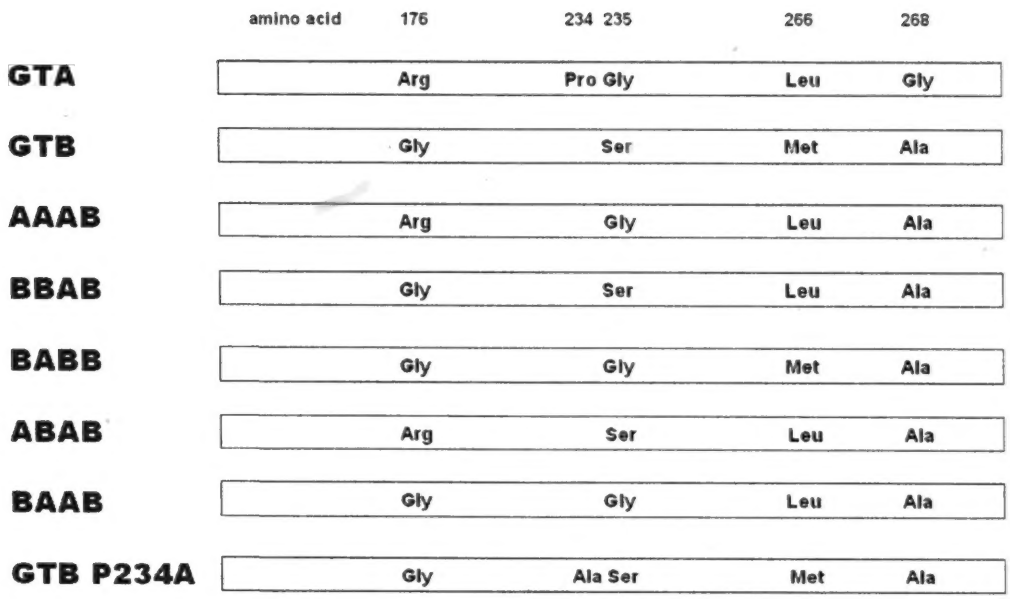

FIG. 1–14. **Schematic representation of the divergent amino acid positions among GTA, GTB and *cis*-AB enzymes.** Only changes from the consensus sequence (GTA) are shown.

1.3 Crystallography

X-ray crystallography is a powerful technique to determine the structures of proteins. Structure-function relationships for a protein can be efficiently deduced by crystallography. The principles of crystallography have been extensively reviewed (32–36); therefore the following section will briefly cover basic crystallography principles and applications to structure determination and analysis used in this thesis.

1.3.1 Crystallization

The first process in protein crystallography is the crystallization of the protein of interest. Various factors can affect the crystallization, for example, pH, temperature, solubility, purity and additives (37). The crystallization method can also influence the crystallization process. Therefore, it is essential to understand the behaviour of the protein and to observe its reaction to various environments. Finding the optimized conditions for

obtaining high quality crystals for diffraction studies may require numerous trials and significant amounts of purified proteins of interest.

A crystal consists of atoms arranged in a pattern that repeats regularly in three dimensions (38). Ideally, individual proteins are similarly affected by the environment and added to the growing crystal during crystallization process. If the rate of crystallization is slow enough to allow each protein molecule to orient itself favourably with an ordered manner, a highly ordered crystal for diffraction studies is formed.

A crystal has a lattice as a set of reference points which can be established where all points have the same chemical environment and are in identical positions with respect to the repeating diffraction pattern. These reference points can be connected to form a lattice that determines a three-dimensional shape of identical building blocks. This building block is called the unit cell. The unit cell is the smallest repeating unit for which its delineating vectors are parallel to important symmetry directions in the lattice. The length of the edges and the angles between the edges determine the size and shape of the unit cell (Fig. 1–15).

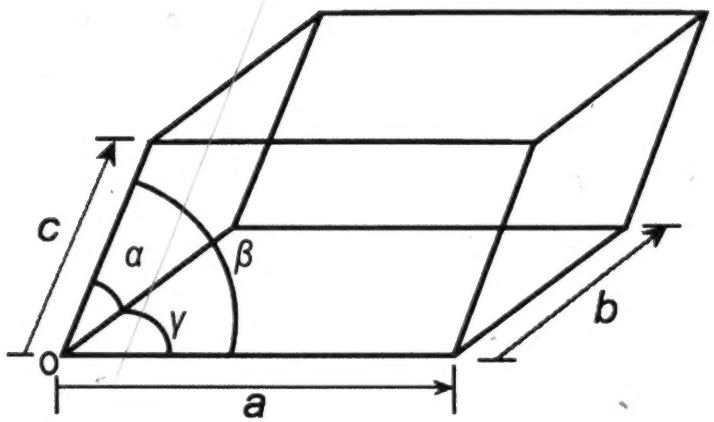

FIG. 1–15. **Schematic representation of the unit cell.** The size and shape of the unit cell are defined by the length of its three edges (a, b, c) and the angles between these edges (α, β, γ) (From Ph. D. thesis of Sonia Patenaude, University of Ottawa 2002 with permission).

1.3.2 X-ray diffraction studies

The wavelength of the radiation must be selected to resolve the features of protein in order to visualize the structure. X-ray diffraction studies are well suited to structural analysis because wavelengths of about 1.0 Å are used. This wavelength is comparable to the ~1.54 Å length of a C–C bond and therefore allows the individual atoms in the protein to be resolved.

Once protein crystals with appropriate size and quality have been crystallized, data collection is carried out. The incident X-ray beam is directed toward the crystal, inducing acceleration of the electrons in the protein and diffraction of X-rays (Fig. 1–16). The scattered X-rays interfere constructively and the resulting diffraction patterns are recorded on a detector. The captured image of these patterns leads to generate electron density maps of protein crystal.

1.3.3 Bragg's law

Crystals can be modeled by separated sections by sets of parallel planes (Fig. 1–17).

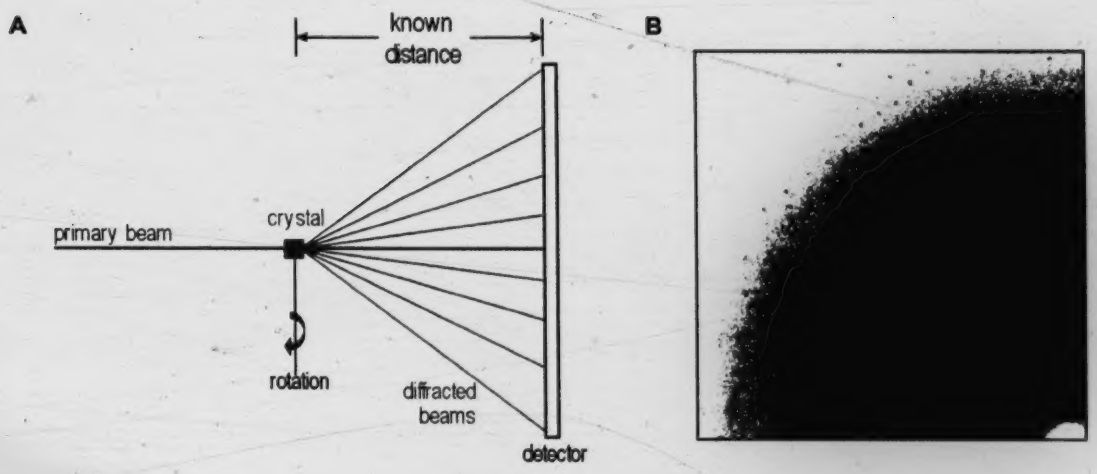

FIG. 1–16. **Schematic representation of X-ray diffraction.** (*A*) A generator produces a primary beam into the protein crystal (*blue*). The interaction of the primary X-rays with the protein's electrons causes a scattering of the X-rays. (*B*) The resulting diffraction patterns are recorded on a detector where they appear as spots. During the collection of diffraction pattern, the crystal is rotated to obtain the maximum number of reflections (From Ph. D. thesis of Sonia Patenaude, University of Ottawa 2002 with permission).

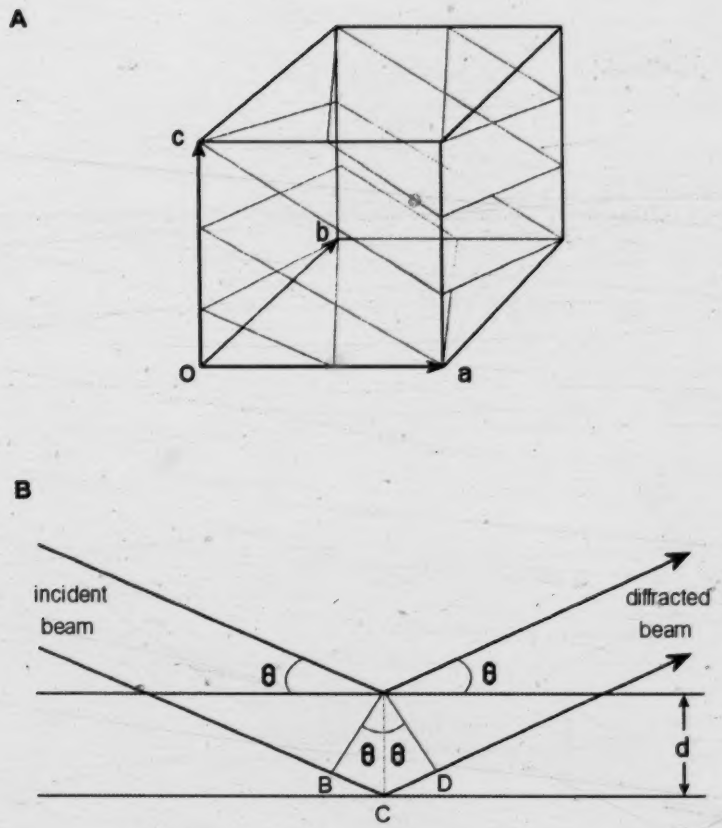

FIG. 1–17. **Mathematical construction of crystal and Bragg's law.** The crystal is constructed as sets of parallel planes (*A*). According to Bragg's law the primary X-ray beam (*red*) strikes adjacent planes, separated by a distance *d*, at an angle θ. The beam is reflected by the same angle (θ). X-rays (*red*) that are reflected from the lower plane have traveled farther than those from the upper plane by a distance BC+CD (*B*) (From Ph. D. thesis of Sonia Patenaude, University of Ottawa 2002 with permission).

Each set of planes is defined by Miller indices *hkl*, where *h* is the number of sections where the *a* axis is cut, *k* is the number of sections where the *b* axis is cut and *l* is the number of sections where the *c* axis is cut. Each set of planes contributes to one observation on the diffracted pattern. Similarly, each observation is given the corresponding Miller indices. Bragg's law suggested that the diffraction of X-rays from crystals is equivalent to their reflection from various sets of planes in the crystal (39) (Fig. 1–17–*B*). The conditions for reflection are satisfied when $\lambda=2d\sin\theta$ where λ is the X-ray wavelength, d is the distance between parallel planes in the crystal and θ is the angle of incidence and diffracted beams of the X-rays.

1.3.4 Structure factor

Every reflection can be described by a structure factor (**F**), which is composed of its amplitude (*F*) and phase (α). The structure factor **F**(*hkl*) for reflection *h*, *k*, *l* is the summation of the contributions of phase and amplitude of each atom in the unit cell (Fig. 1–18–*A*).

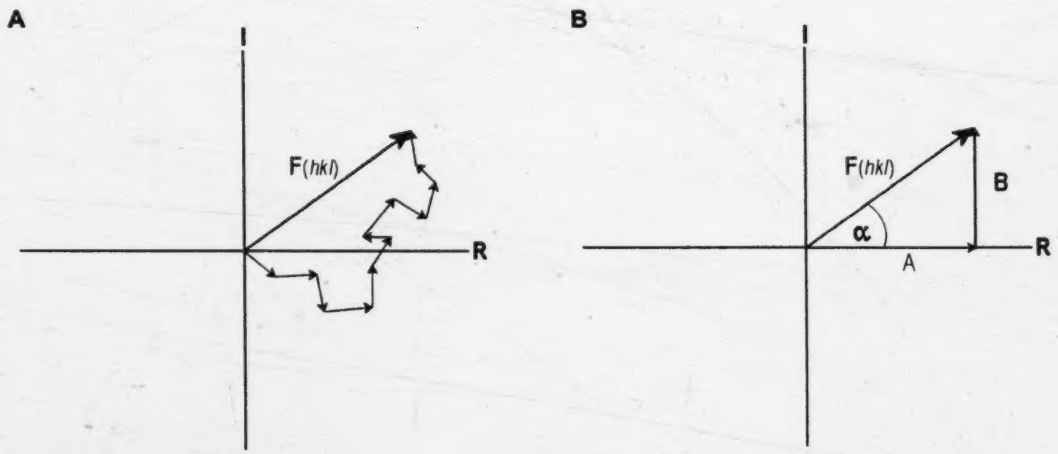

FIG. 1–18. **Vector diagram of structure factor.** The amplitude F and phase α of a reflection can be described by a structure factor. In 1–18–*A*, the contribution of individual atoms (*red*) to a resulting reflection are summed to give the structure factor (*blue*) of that reflection. The Argand diagram in 1–18–*B* gives the structure factor (*blue*) in terms of its real (A, *green*) and imaginary (B, *red*) components (From Ph. D. thesis of Sonia Patenaude, University of Ottawa 2002 with permission).

where N is the total number of atoms and x, y, z denote the fractional coordinates of atom j.

$$\mathbf{F}(hkl) = F(hkl)\cos\{\alpha(hkl)\} + iF(hkl)\sin\{\alpha(hkl)\}$$

$$= A + iB \qquad \text{(Eq. 1–2)}$$

Eq. 1–2 describes the structure factor in terms of its real (*A*) and imaginary (*iB*) parts, which are represented on the Argand diagram (Fig. 1–18–*B*).

In the crystal the geometrical arrangement of planes allows obtaining equal but opposite reflections. During the rotation of the crystal in the X-ray beam, various planes come into diffracting position. The set of planes reflects the incident X-ray (Fig. 1–19–*A*) and the 180° rotation allows the reflection of X-rays from the opposite side of the same set of planes. If the wavelength is distant from absorption edges of the atoms in the crystal, the resulting intensities of both reflections will be equal, but their phases will be opposite; this is known as the Friedel's law (Fig. 1–19–*B*).

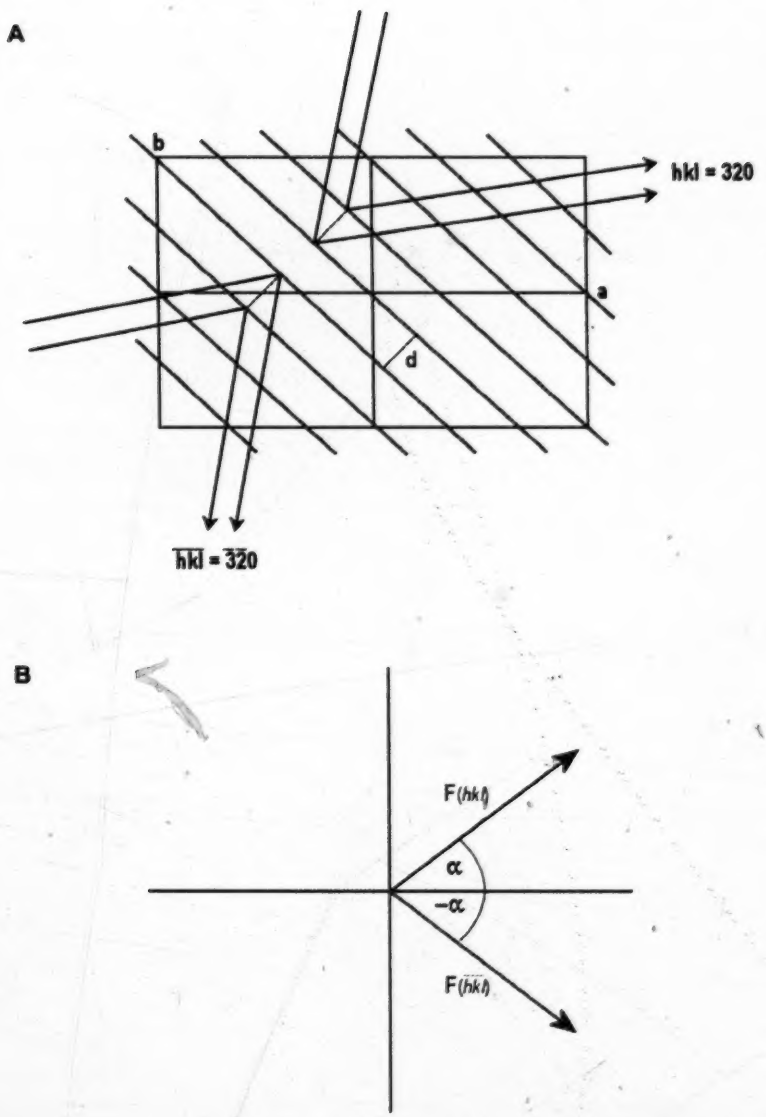

FIG. 1–19. **Schematic representation of Friedel's law.** The geometrical arrangement of Bragg planes (*blue*) given in (*A*) allows the X-rays (*red*) to be reflected from opposite sides of the same set of planes on a 180 ° rotation of the crystal. Friedel's law (*B*) states that the intensities of the resulting reflections will be equal ($\mathbf{F}(hkl) = \mathbf{F}(\overline{hkl})$), but their phases will be opposite ($\alpha(hkl) = -\alpha(\overline{hkl})$) (From Ph. D. thesis of Sonia Patenaude, University of Ottawa 2002 with permission).

1.3.5 Scattered factor

The X-ray scattering of an atom for a reflection is strongly dependent upon the number of electrons in that atom and is known as the scattering factor (f_o). It has been described that the atomic scattering factor is influenced by Bragg angle and thermal motion. Therefore, the revised scattering factor (f) is given by Eq. 1–3.

$$f = f_o e^{-B(\sin^2\theta)/\lambda^2} \qquad \text{(Eq. 1–3)}$$

where B is the correction for thermal motion and $(\sin^2\theta) / \lambda^2$ is the correction term for scattering angle and wavelength. Thermal motion is often called the B-factor and it is related to the mean square displacement (u^2) of a given atom by $B = 8\pi^2 u^2$

1.3.6 Phase problem

An electron density distribution can be calculated when the structure factors and phases are determined. Unfortunately, while the structure factors of the diffraction pattern

can be calculated, all information on phases is lost. Therefore, it is required to find the missing information that is the source of phase problem. Various methods can be used to solve the phase problem in crystallography such as multiple isomorphous replacement (MIR) and molecular replacement (MR).

1.3.6.1. Patterson function

The Patterson function (Eq. 1–4) can be used as a tool for solving the phase problem. It does not need phase information and uses only the easily accessible intensities (amplitudes squared) of each reflection. Because phase information is not used in the Patterson function, the result cannot be interpreted as a set of atomic positions, but as a collection of interatomic vectors.

$$P(x,y,z) = \frac{1}{V}\sum_{h}\sum_{k}\sum_{l}\left|\mathbf{F}_{hkl}\right|^2 \cos 2\pi(hx+ky+lz) \qquad \text{(Eq. 1–4)}$$

where x, y, and z are atomic coordinates. The Patterson map calculated with $|\mathbf{F}|^2$ gave peaks corresponding to all of the interatomic vectors. The information obtained from Patterson

maps is useful for solving crystal structures by various methods including MIR, MR.

1.3.6.2. Multiple isomorphous replacement (MIR)

In multiple isomorphous replacement (MIR), heavy atoms are introduced into the protein crystal, most commonly by soaking. These atoms provide observable changes in the diffraction pattern of the protein crystal. The position of the heavy atoms can be calculated by comparing the intensities of diffraction patterns of the native and altered protein crystal and the phase problem resolved. The general principle of MIR follows.

The structure factor for the altered crystal ($\mathbf{F}_{PH}$) means

$$\mathbf{F}_{PH} = \mathbf{F}_P + \mathbf{F}_H \qquad \text{(Eq. 1–5)}$$

where $\mathbf{F}_P$ is the structure factor of native protein and $\mathbf{F}_H$ is structure factor of the heavy atoms (Fig. 1–20). Therefore, if $\mathbf{F}_P$ and $\mathbf{F}_{PH}$ are measured, $\mathbf{F}_H$ can be calculated. Using the value $\mathbf{F}_H$, a Patterson map can be computed from which the positions of the heavy atoms

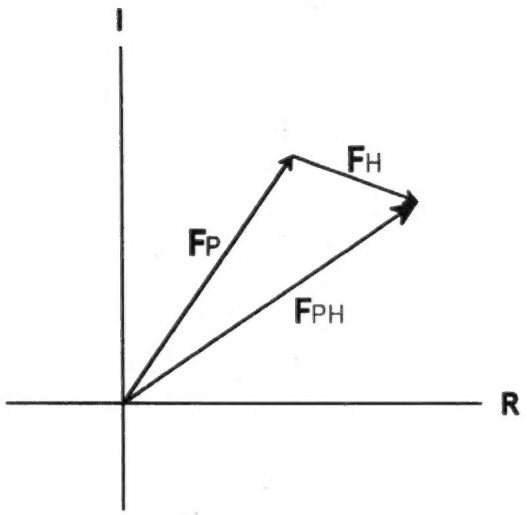

FIG. 1–20. **Contributions of vectors to the structure factor in MIR.** The contributions of the native protein ($\mathbf{F_P}$, *black*) and the heavy atom ($\mathbf{F_H}$, *red*) to the observed structure factor ($\mathbf{F_{PH}}$, *blue*) are represented (From Ph. D. thesis of Sonia Patenaude, University of Ottawa 2002 with permission).

can be calculated and phase angle α_H can be determined. Fig. 1–21–*A* shows how the $\mathbf{F}_{PH}$, $\mathbf{F}_H$ and α_H can be combined to determine the phase of $\mathbf{F}_P$. However, the solution has two possibilities that make undetermined value of α_P. To resolve this uncertainty, it is required to use a second heavy atom derivative (Fig. 1–21–*B*). This process can determine a single value for α_P.

1.3.6.3. Molecular replacement (MR)

Molecular replacement is an alternative approach for solving the phase problem. This method determines the phases by a model. In the molecular replacement approach, it is important that the structural model is highly homologous to the protein. Sequence identity is thought to be correlated to the level of similarity of two structures. The higher the similarity between the model and the protein the easier that the phase problem will be solved by a MR solution. In contrast if a very poor model is used to solve the phase problem, even if MR succeeds, the phase information is unreliable to successfully refine the structure due to serious model bias.

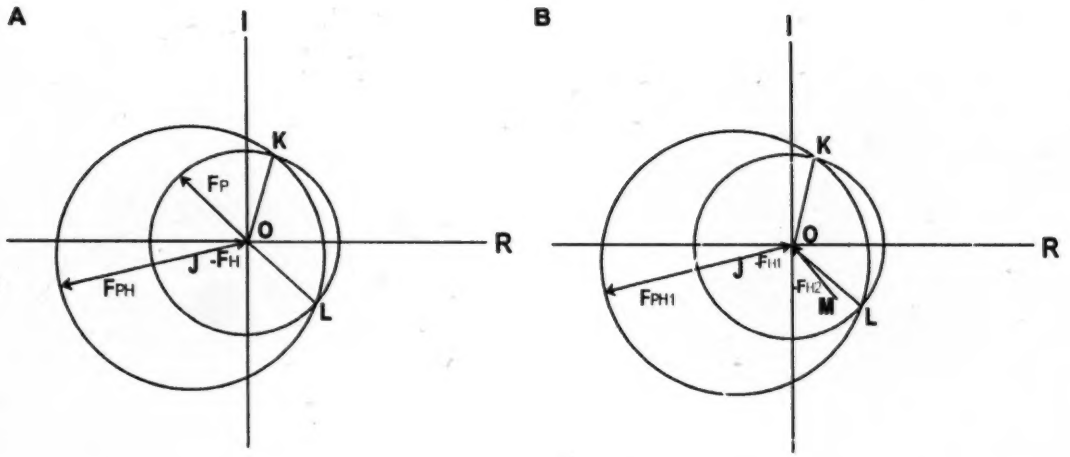

FIG. 1–21. **Schematic diagram of the phase calculation using MIR**. The single isomorphous replacement situation shown in (*A*). Subtraction of the heavy atom contribution ($\mathbf{F_H}$, *red*) from the origin vector $\mathbf{F_{PH}}$ (*green*) at J can generate the contribution of protein crystal ($\mathbf{F_P}$, *blue*). A circle of radius $\mathbf{F_P}$, centered about O, makes two intersections and therefore two possibilities of $\mathbf{F_P}$ (OK and OL). This uncertainty can be corrected by MIR, represented in (*B*). The contribution of a second heavy atom derivative ($\mathbf{F_{PH2}}$, *yellow*) is drawn with its center at M, making the unique solution of $\mathbf{F_P}$ (OK) (From Ph. D. thesis of Sonia Patenaude, University of Ottawa 2002 with permission).

The purpose of the MR approach is to orient and position the model that coincides with the position of the protein in the crystal. The rotation function exploits the intramolecular vectors of the Patterson map, which rely on the orientation of the molecule in the unit cell. The translation function makes use of intermolecular vectors, which depend on the orientation and position of the molecule in the unit cell. The resulting rotation angles and three-dimensional translation place the model. The next step is to transform to search model into the protein by inserting the correct amino acid sequence.

1.4. Overview of thesis

The goal of my thesis was to investigate the kinetic behaviour of blood group O^2 and *cis*-AB glycosyltransferases. This required the preparation of mutant enzymes related to O^2 and *cis*-AB enzymes by site-directed mutagenesis, their expression in *E. coli* and purification to homogeneity using ion-exchange and affinity chromatography. Then kinetic characterization of purified enzymes was carried out and X-ray diffraction studies were performed to reveal structural features of the mutants.

References

1. Landsteiner, K. (1900) *Zentbl. Bakt. Parasitkde.* **27,** 357–366

2. Landsteiner, K. (1901) *Blutes. Wein. Klin. Wschr.* **14,** 1132–1134

3. Watkins, W. M. (1980) *Adv. Hum. Genet.* **10,** 379–385

4. Palcic, M. M., Seto, N. O., and Hindsgaul, O. (2001) *Transfus. Med.* **11,** 315–323

5. Gastinel, L. N., Bignon, C., Misra, A. K., Hindsgaul, O., Shaper, J. H., and Joziasse, D. H. (2001) *EMBO J.* **220,** 638–649

6. Yamamoto, F., Clausen, H., White, T., Marken, J., and Hakomori, S. (1990) *Nature* **345,** 229–233

7. Paulson, J. C., Weinstein, J., Ujita, E. L., Riggs, K. J., and Lai, P. H. (1987) *Biochem. Soc. Trans.* **15,** 618–620

8. Wickner, W. T., and Lodish, H. F. (1985) *Science* **230,** 400–407

9. Seto, N. O. L., Palcic, M. M., Hindsgaul, O., Bundle, D. R., and Narang, S. A. (1995) *Eur. J. Biochem.* **234**, 323–328

10. Seto, N. O. L, Palcic, M. M., Compston, C. A., Li, H., Bundle, D. R., and Narang, S. A. (1997) *J. Biol. Chem.* **272,** 14133–14138

11. Patenaude, S. I., Seto, N. O. L., Borisova, S. N., Szpacenko, A., Marcus, S. L., Palcic, M. M., and Evans, S. V. (2002) *Nat. Strut. Biol.* **9,** 685–690

12. Breton, C., Bettler, E., Joziasse, D. H., Geremia, R. A., and Imberty, A. J. (1998) *J. Biochem.* **123,** 1000–1009

13. Busch, C., Hofmann F., Selzer, J. Munro, S., Jeckel, D., and Aktories, K. (1998) *J. Biol. Chem.* **273,** 19566–19672

14. Seto, N. O. L., Compston, C. A., Evans, S. V., Bundle, D. R., Narang, S. A., and Palcic, M. M. (1999) *Eur. J. Biochem.* **259,** 770–775

15. Evans, S. V. (1993*) J. Mol. Graphics.* **11,** 134–138

16. Bennett, E. P., Steffensen, R., Clausen, H., Weghuis, D. O., Vankessel, A. G. (1995) *Biochem. Biophys. Res. Commun.* **206,** 318–325

17. Yamamoto, F., McNeill, P. D., and Hakomori, S. (1995) *Glycobiology* **5,** 51–58

18. Yamamoto, F., Marken, J., Tsuji, T., White, T., Clausen, H., and Hakomori, S. (1990) *J. Biol. Chem.* **265,** 1146–1151

19. Yamamoto, F., and Hakomori, S. (1990) *J. Biol. Chem.* **265,** 19257–19262

20. Olsson, M. L., Thuresson, B., and Chester, M. A. (1995) *Biochem. Biophys. Res. Commun.* **216,** 642–647

21. Yamamoto, F., McNeill, P. D., Yamamoto, M., Hakomori, S., Bromilow, I. M., and Duguid, J. K. (1993) *Vox Sang.* **64,** 175–178

22. Yamamoto, F., and McNeill, P. D. (1996) *J. Biol. Chem.* **271,** 10515–10520

23. Amado, M., Bennett, E. P., Carneiro, F., and Clausen, H. (2000) *Vox Sang.* **79,** 219–226

24. Olsson, M. L., and Chester, M. A. (1996) *Vox Sang.* **71,** 113–117

25. Olsson, M. L., and Chester, M. A (2001) *Transfus. Med.* **11,** 295–313

26. Ogasawara, K., Yabe, R., Uchikawa, M., Saitou, N., Bannai, M., Nakata, K., Takenaka, M., Fujisawa, K., Ishikawa, Y., Juji, T., and Tokunaga, K. (1996) *Blood* **88,** 2732–2737

27. Yamamoto, F., McNeill, P. D., Kominato, Y., Yamamoto, M., Hakomori, S., Ishimoto, S., Nishida, S., Shima, M., and Fujimura, Y. (1993) *Vox Sang.* **64,** 120–123

28. Fukumori, Y., Ohnoki, S., Yoshimura, K., Nagao, N., Shibata, H., Tomita, T., Okubo, Y., and Yamaguchi, H. (1996) *Transfus. Med.* **6,** 337–344

29. Mifsud, N. A., Watt, J. M., Condon, J. A., Haddad, A. P., and Sparrow, R. L. (2000) *Transfusion* **40,** 1276–1277

30. Yu, Lung-Chih, Lee, Hui-Lin, Chan, Yung-Shi, and Lin, M., (1999) *Biochem. Biophys. Res. Commun.* **262,** 487–493

31. Yamamoto, M., Lin, X., Kominato, Y., Hata, Y., Noda, R., Saitou, N., and Yamamoto, F. (2001) *J. Biol. Chem.* **276,** 13701–13708

32. Sands, D. E. (1969) *Introduction to crystallography*, Dover Publications, Inc., Mineola, N.Y, p.165

33. Glusker, J. P., and Trueblood, K. N. (1972) *Crystal structure analysis: a primer*, Oxford University Press, Oxford, p.192

34. Blundell, T. L., and Johnson, L. N. (1976) *Protein crystallography*, Academic Press, New York, p.565

35. Stout, G. H., and Jensen, L. H. (1989) *X-ray structure determination: a practical guide*. 2nd Ed., Wiley, New York, p.453

36. Woolfson, M. M. (1997) *An introduction to X-ray crystallography*. 2nd Ed., Cambridge University Press, New York, NY, USA, p.402

37. McPherson, A. (1999) *Crystallization of biological macromolecules*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, p.586

38. Barrett, C. S. (1952) *Structure of metals, crystallographic methods, principles, and*

data. 2nd Ed., McGraw-Hill, New York, p.661

39. Bragg, W. L. (1913) *Proc. Camb. Phil. Soc.* **17**, 43–57.

Chapter 2. Characterization of Recombinant Human Blood Group O^2 Glycosyltransferase

2.1 Introduction

The human ABO blood group glycosyltransferases, which synthesize the A and B antigens, catalyze the transfer of a monosaccharide residue from UDP-GalNAc or UDP-Gal to Fucα1–2Galβ-R (H) acceptors (1, 2). The A synthesizing α1–3 *N*-acetylgalactosaminyltransferase (GTA, EC 2.4.1.40) transfers GalNAc from UDP-GalNAc to H terminating acceptors, producing the A antigen GalNAcα1–3[Fucα1–2]Galβ-R. The B synthesizing α1–3 galactosyltransferase (GTB, EC 2.4.1.37) transfers Gal from UDP-Gal to the same acceptor, producing the B antigen Galα1–3[Fucα1–2]Galβ-R (Fig. 2–1). Blood group O enzymes do not synthesize either antigen, and *cis*-AB enzymes synthesize both (3). GTA and GTB are highly homologous enzymes differing in only four of 354 amino acids, Arg-176 → Gly, Gly-235 → Ser, Leu-266 → Met, and Gly-268 → Ala (Fig. 2–2) (3). Alteration of the four critical amino acids converts the specificity from GTA to GTB. X-ray diffraction studies for GTA and GTB

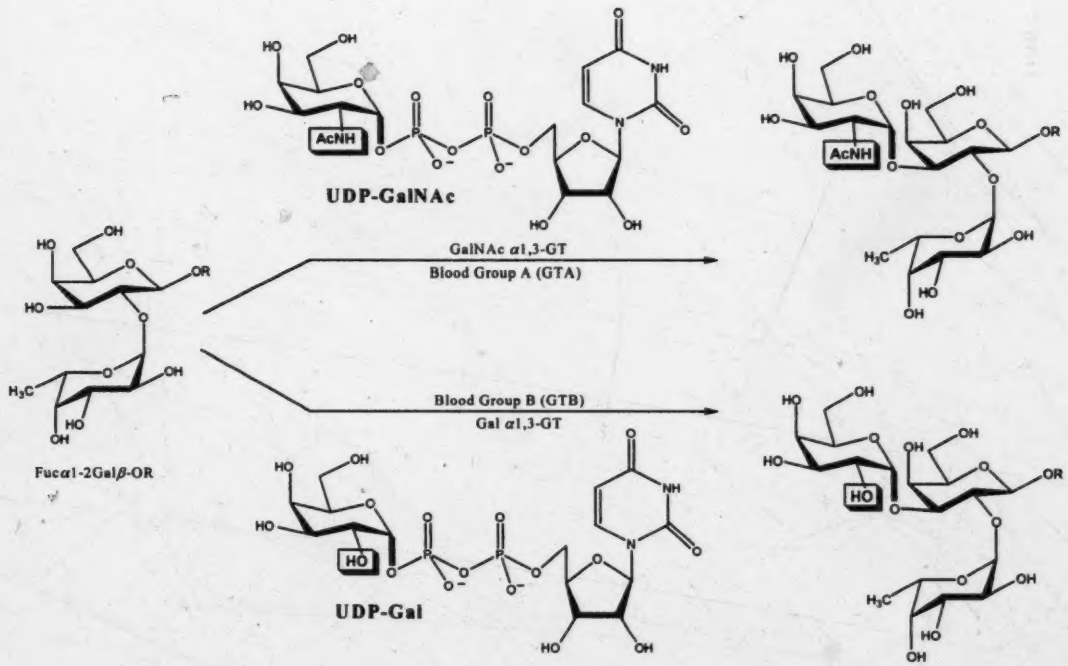

FIG. 2–1. **The reactions catalyzed by the human blood group A and B glycosyltransferases GTA and GTB.**

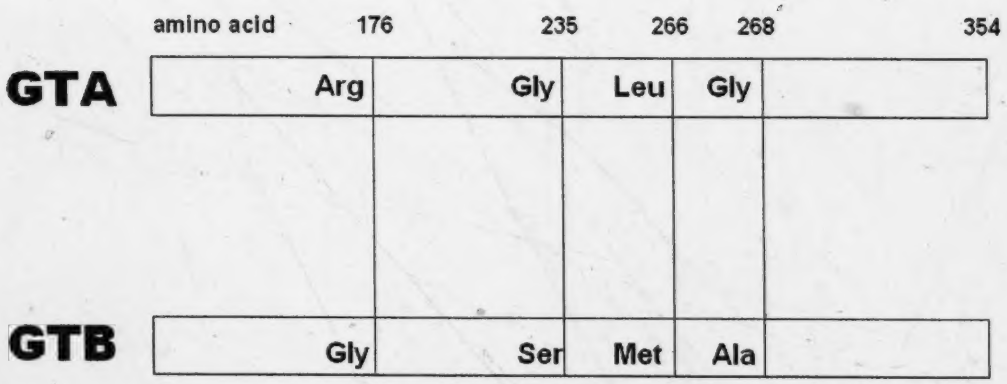

FIG 2–2. **Schematic diagram of the divergent amino acid positions for GTA and GTB.** Critical amino acid residues are shown. Alteration of these critical residues determines the specificity of donor utilized.

complexes with H acceptor and UDP have shown the structural basis for substrate binding as well as the significant contacts (Fig. 2–3) (4, 5).

The origin of blood group O glycosyltransferases was originally shown to be a deletion or mutation in the GTA or GTB gene that gave inactive truncated enzyme (3). More recently, an O^2 enzyme was discovered that was a full-length form of GTA with proline 74 replaced with serine, arginine 176 replaced with glycine, and glycine 268 replaced with arginine (Fig. 2–4) (6, 7). The change of the glycine to arginine at codon 268 adversely affects nucleotide sugar binding such that the enzyme is inactive (8, 9).

In previous studies, the O^2 glycosyltransferase showed no measurable transferase activity (7, 10, 11). Greenwell and Watkins (12) showed that pooled and highly concentrated group O plasma appeared to contain A^2-like transferase activity; however, no kinetic studies were performed because of the extremely low activity of this enzyme. Here we prepare three mutant enzymes where proline 74 was replaced with serine, glycine 268 was replaced with arginine, and with both residues at 74 and 268 replaced with serine and arginine. The mutants were based on BAAA, where this four-letter code denotes whether the critical amino acid is in GTA or GTB. In this code GTA is designated AAAA (Arg-176, Gly-235, Leu-266, Gly-268), GTB is BBBB (Gly-176, Ser-235,

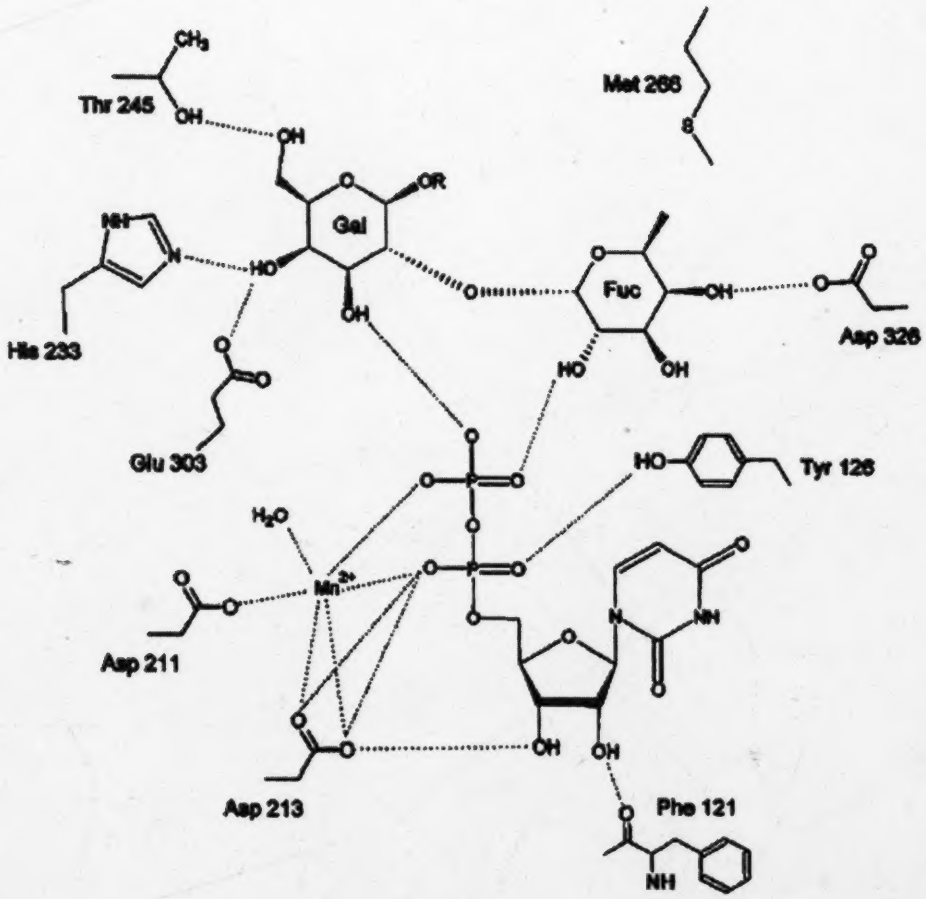

FIG. 2–3. **Schematic representation of significant contacts observed between GTB, acceptor, and UDP.** (From Patenaude, S. I., Seto, N. O. L., Borisova, S. N., Szpacenko, A., Marcus, S. L., Palcic, M. M., and Evans, S. V. (2002) *Nat. Struct. Biol.* **9**, 685-690).

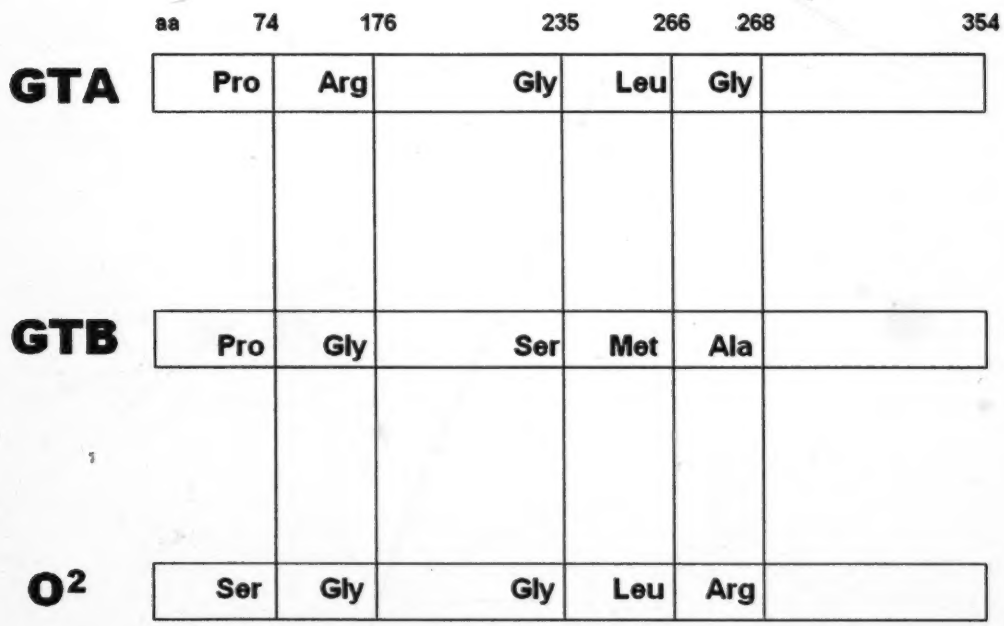

FIG. 2–4. **Schematic diagram of the amino acid positions among GTA, GTB and O^2 glycosyltransferase.** Amino acid positions are represented by lines.

Met-266, Ala-268) and BAAA is Gly-176, Gly-235, Leu-266, Gly-268. We use site-directed mutagenesis followed by purification, kinetic characterization and X-ray crystallography to determine which mutation has a dominant effect on activity.

2.2 Experimental procedures

2.2.1 Materials and general techniques

All molecular biology procedures were carried out according to standard procedures (13, 14). Reagents and supplies are tryptone (Bacto); yeast extract and casamino acids (Difco); glycerol (GibcoBRL); Na_2HPO_4, dithiothreitol (DTT), 3-(*N*-morpholino) propanesulfonic acid (MOPS) (Fisher); isopropyl-1-thio-β-D-galactopyranoside (IPTG) (Rose Scientific); protease inhibitor cocktail tablets (Complete, EDTA-free – Roche); PCR reagents including *Pfx* DNA polymerase, *Pfx* buffer, and DNA ladder (Invitrogen); restriction enzymes and DNA-modifying enzymes (New England Biolabs and Life Technologies, Inc.); primers including HJL01, HJL02, HJL03 (IDT); BL21-gold competent cells (Stratagene); DYEnamic ET terminator cycle sequencing kit

(Amersham Biosciences); Mini-sub Gel GT for Agarose gel electrophoresis for PCR reaction and protein assay dye reagent concentrate (Bio-Rad); KH_2PO_4, NH_4Cl, NaCl, $MnCl_2$, $CaCl_2$, thiamine hydrochloride (vitamin B1), D[-]-α-aminobenzylpenicillin (ampicillin), ethylenediaminetetraacetic acid (EDTA-tetrasodium salt: hydrate), glucose, UDP, UDP-GalNAc, and UDP-Gal, GenElute Plasmid Miniprep Kit, Post-Reaction Purification Columns (Sigma); QIAquick PCR Purificatioin Kit, QIAquick Gel Extraction Kit (Qiagen); the media for SP-Sepharose Fast Flow and UDP-hexanolamine columns (Pharmacia Fine Chemicals); the radioactive [6-^{3}H]-labeled analogs of UDP-GalNAc and UDP-Gal (American Radiolabeled Chemicals); Sep-Pak C_{18} (+) reverse phase cartridges (Waters); Ecolite (+) liquid scintillation cocktail (ICN); Centriplus-10 protein concentrators (Amicon); bovine gamma globulin (IgG), Slide-a-Lyzer dialysis cassettes (Pierce Chemical Co.); SDS-PAGE system (Hoefer Scientific Instruments) were purchased commercially. The acceptor substrate Fucα1–2Galβ-O$(CH_2)_7CH_3$ was a gift from Dr O. Hindsgaul (University of Alberta). A Rigaku R-AXIS4++ detector, MicroMax X-ray generation system (Rigaku/MSC) and Osmic "Blue" confocal x-ray mirrors (Osmic) were used for the crystallography.

2.2.2 Cloning of P74S

The original GTB and GTA gene sequences (aa 54–354) were described in previous papers (15, 16). In this chapter, GTB (aa 63–354) and GTA (aa 63–354) are denoted as wild type enzymes. The BAAA P74S mutant (Ser-74, Gly-176, Gly-235, Leu-266, Gly-268) was constructed by PCR using the BAAA clone as a template. The forward primer HJL01 (5'-A TAT GAA TTC ATG GTT TCC CTG CCG CGT ATG GTT TAC CCG CAG TCC AAA GTT CTG ACC CCA TGC CG-3') introduced an *Eco*RI site at the 5' end, and the reverse primer PCR3B (5'-ATA ATT AAG CTT CTA TCA CGG GTT ACG AAC AGC CTG GTG GTT TTT-3') introduced a *Hin*dIII site at the 3' end of the BAAA gene. The amplified genes were digested by restriction enzymes (*Eco*RI, *Hin*dIII) and ligated into the previously digested pCWΔlac vector (17). The ligation reaction was incubated at room temperature overnight and transformed into *Escherichia coli* BL21-gold competent cells. A single transformant was inoculated into LB broth containing ampicillin and incubated overnight at 37 °C. Plasmids were purified with a mini plasmid preparation column. The entire sequence was confirmed by sequencing using a DYEnamic ET terminator cycle sequencing kit.

2.2.3 Cloning of G268R

The BAAA G268R mutant was constructed by PCR using the BAAA clone as a template. The forward primer MIN2 together with HJL02 (5'-ACC GAA GAA ACG ACC CAG GTA GTA GAA GTC ACC-3') introduced an *Eco*RI site at the 5' end, and the reverse primer PCR3B together with HJL03 (5'-C CTG GGT CGT TTC TTC GGT GGT TCC GTT CAG-3') introduced a *Hin*dIII site at the 3' end of the gene. The two overlapping fragments were isolated and amplified using the outside primers MIN2 and PCR3B. The amplified gene was digested with restriction enzymes (*Eco*RI, *Hin*dIII) and treated as described for the cloning of P74S.

2.2.4 Cloning of P74S & G268R

The BAAA P74S & G268R mutant was constructed by PCR using the BAAA G268R clone as a template. The forward primer HJL01 (5'-A TAT GAA TTC ATG GTT TCC CTG CCG CGT ATG GTT TAC CCG CAG TCC AAA GTT CTG ACC CCA TGC CG-3') introduced an *Eco*RI site at the 5' end, and the reverse primer PCR3B (5'-ATA

ATT AAG CTT CTA TCA CGG GTT ACG AAC AGC CTG GTG GTT TTT-3') introduced a *Hin*dIII site at the 3' end of the gene. The amplified gene was digested with restriction enzymes (*Eco*RI, *Hin*dIII) and treated as described for the cloning of P74S.

2.2.5 Enzyme purification

E.coli cells were grown overnight, at 30 °C, with shaking at 200 rpm, in 500 ml baffled flasks, in a starter culture containing 1 × TB broth (13.2 g of tryptone, 26.4 g of yeast extract, and 4.4 ml of glycerol/L), 1 × M9 salts (6.6 g of Na_2HPO_4, 3.3 g of KH_2PO_4, 1.1 g of NH_4Cl, and 0.55 g of NaCl/L), and M9 supplements (1 mM $MgCl_2$, 0.1 mM $CaCl_2$, 5 μg/ml vitamin B1, 0.2% glucose, 0.4% casamino acids, and 0.1 mg/ml ampicillin/L). The cultures were scaled up to 4 × 1 liter size (each 1 liter culture was in a 4 liter baffled flask) by using 30–100 ml of starter culture per 1 liter of scale-up broth (with the same concentration of all the materials as that of the starter broth), and grown with shaking at 200 rpm, 30 °C. The cultures were induced with 1 mM IPTG at OD_{600} 0.6–1.0 and harvested 20–24 h later. The cells were then spun down at 11,325 × *g* for 20 min, at 4 °C, to remove the supernatant. The pelleted cells were then frozen at -20 °C for

purification later or resuspended immediately in 200 ml of 50 mM MOPS, 1 mM DTT, pH 7.0, on ice. A protease inhibitor cocktail tablet was added to the cell suspension, which was then passed once through a French Press at 20,000 psi and centrifuged at 105,000 × *g* for 60 min at 4 °C. The resulting supernatant was then loaded onto an SP-Sepharose FF column (410 ml, 5.8 cm × 15.5 cm) at a flow rate of 5 ml/min The SP-Sepharose FF column was regenerated between enzyme purifications by washing with 8 M urea, 2 M NaCl, 1 M NaOH and 30% (v/v) of isopropanol. The resin was assayed for activity to ensure that no contaminating enzymes were present. The column was equilibrated with 50 mM MOPS, 1 mM DTT, pH 7.0 prior to loading of the supernatant from ultracentrifugation. After loading, the column was washed with 50 mM MOPS, 1 mM DTT, pH 7.0, at 5 ml/min until OD_{280} was at the baseline (approximately 3 column volumes). The enzyme was eluted with 50 mM MOPS, 1 mM DTT, 0.5 M NaCl, pH 7.0, at 4 ml/min collecting 12 ml fractions. The fractions with appreciable amounts of enzymatic activity, as assessed by a standard assay for GTA activity (described below), were then pooled for loading onto a UDP-Hex column. Manganese chloride was added to the pooled fractions to a final concentration of 5 mM prior to loading onto the UDP-Hex column (25 ml, 1.8 cm × 13.1 cm). The UDP-Hex column was also regenerated by

washing with 8 M urea between consecutive enzyme purifications to avoid contamination, and equilibrated with 50 mM MOPS, 1 mM DTT, 0.5 M NaCl, 5 mM $MnCl_2$, pH 7.0, at a flow rate of 1–2 ml/min, 4 °C. Loading was done at a flow rate of 0.2–0.4 ml/min, overnight, followed by a wash with equilibration buffer at a flow rate of 0.6–0.8 ml/min until OD_{280} was at the baseline. The enzyme was eluted with the loading buffer containing 10 mM UDP at a flow rate of 1 ml/min collecting 4 ml fractions. The fractions were assayed for GTA activity, and the active fractions were pooled, transferred into Slide-a-Lyzer dialysis cassettes and dialyzed against 2 × 4 liters of 50 mM MOPS, 1 mM DTT, 100 mM NaCl, 5 mM $MnCl_2$, pH 7.0, to remove the UDP. The samples that had a low activity were concentrated in a centrifuge using a Centriplus 30 filtration unit. After dialysis, the samples were assayed for activity and protein concentration by using the saturating assay and mass quantitation methods described below. SDS-PAGE was used to confirm homogeneity as described (15).

2.2.6 Standard assays for GTA activity

To assess the activity of an enzyme during purification, the following standard

assays were employed.

2.2.6.1 P74S mutant standard assay

The enzyme was diluted in a 2 × A/B buffer (2 mg/ml BSA, 40 mM $MnCl_2$, 100 mM MOPS, pH 7.0), and 5 μl of this were added to 5 μl of GTA standard assay mix (140 μM Fucα1–2Galβ-*O*-$(CH_2)_7CH_3$ acceptor, 140 μM UDP-GalNAc donor, and 0.02 μCi UDP-[6-^{3}H]GalNAc donor (approximately 44,000 dpm)). The incubation time was 30 min, at 37 °C, and the amounts of acceptor and donor consumed were less then 10% of the MAX. At the end of the incubation, the reaction was quenched with 0.5 ml cold deionized water, and the reaction product was transferred to a Sep-Pak C_{18} reverse phase cartridge. The cartridge was previously washed with 50 ml MeOH then 50 ml deionized water. After sample loading the cartridge was washed with 50 ml deionized water, and the radiolabeled product was eluted with 3.5 ml MeOH into a scintillation vial. Liquid scintillation cocktail (10 ml) was added to each vial for counting (5 min) in a Beckman liquid scintillation counter (LS 1801).

2.2.6.2 G268R mutant standard assay

Undiluted enzyme (5 μl) was added to 5 μl of GTA assay mix (140 μM Fucα1–2Galβ-O-$(CH_2)_7CH_3$ acceptor, 140 μM UDP-GalNAc donor, 2 × A/B buffer, and 0.08 μCi UDP-[6-^{3}H]GalNAc donor (approximately 178,000 dpm)). The incubation time was 4 h, at 37 °C, and the amounts of acceptor and donor consumed were less then 10% of the MAX. At the end of the incubation time, the reaction was quenched with 0.5 ml cold deionized water and was treated as described for the P74S mutant standard assay.

2.2.6.3 P74S & G268R mutant standard assay

Undiluted enzyme (5 μl) was added to 5 μl of GTA assay mix (140 μM Fucα(1–2)Galβ-O-$(CH_2)_7CH_3$ acceptor, 140 μM UDP-GalNAc donor, 2 × A/B buffer, and 0.03 μCi UDP-[6-^{3}H]GalNAc donor (approximately 66,600 dpm)). The incubation time was 4 h, at 37 °C, and the amounts of acceptor and donor consumed were less then 10% of the MAX. At the end of the incubation time, the reaction was quenched with 0.5 ml cold deionized water, and was treated as described for the P74S mutant standard assay.

2.2.7 Saturating assay for GTA activity

The standard assay mix was modified and contained sufficient acceptor and donor to fully saturate the enzymes. These were used to assess the final activity of the respective enzymes.

2.2.7.1 P74S mutant saturating assay

The enzyme was diluted in a 2× A/B buffer (2 mg/ml BSA, 40 mM $MnCl_2$, 100 mM MOPS, pH 7.0), and 5 μl of this was added to 5 μl of GTA assay mix (600 μM Fucα1–2Galβ-O-$(CH_2)_7CH_3$ acceptor, 1000 μM UDP-GalNAc donor, and 0.1 μCi UDP-[6-^{3}H]GalNAc donor (approximately 222,000 dpm)). The incubation time was 30 min, at 37 °C, and the amounts of acceptor and donor consumed were less then 10% of the MAX. At the end of the incubation time, the reaction was quenched with 0.5 ml cold deionized water, and was treated as described for the P74S mutant standard assay.

2.2.7.2 G268R mutant saturating assay

Undiluted enzyme (5 μl) was added to 5 μl of GTA assay mix (1200 μM Fucα(1–2)Galβ-O-$(CH_2)_7CH_3$ acceptor, 5000 μM UDP-GalNAc donor, 2 × A/B buffer, and 0.1 μCi UDP-[6-^{3}H]GalNAc donor (approximately 222,000 dpm)). The incubation time was 4 h, at 37 °C, and the amounts of acceptor and donor consumed were less then 10% of the MAX. At the end of the incubation time the reaction was quenched with 0.5 ml cold deionized water, and was treated as described for the P74S mutant standard assay.

2.2.7.3 P74S & G268R mutant saturating assay

Undiluted enzyme (5 μl) was added to 5 μl of GTA assay mix (6000 μM Fucα(1–2)Galβ-O-$(CH_2)_7CH_3$ acceptor, 6000 μM UDP-GalNAc donor, 2× A/B buffer, and 0.2 μCi UDP-[6-^{3}H]GalNAc donor (approximately 444,000 dpm)). The incubation time was 4 h, at 37 °C, and the amounts of acceptor and donor consumed were less then 10% of the MAX. At the end of the incubation time, the reaction was quenched with 0.5 ml cold deionized water, and was treated as described for the P74S mutant standard assay.

2.2.8 Protein determination

The Bio-Rad protein assay procedure based on Bradford's method (18) was used to estimate protein concentrations. Bovine gamma globulin was used as a protein standard at 0.05, 0.1, 0.2, 0.4, 0.6 mg/ml.

2.2.9 Kinetic characterization

Assays were carried out at 37 °C in a total volume of 15 μl containing substrates and enzyme in 50 mM MOPS buffer, pH 7.0, with 20 mM $MnCl_2$ and 1 mg/ml BSA. Seven different concentrations of the donor or acceptor were used, and the amount of substrate consumed was less than 15% to ensure linear initial reaction rates. Data were analyzed by fitting to the Michaelis-Menten equation (Eq. 1).

$$\upsilon = \frac{V_{\max}[S]}{K_m + [S]} \qquad \text{(Eq. 1)}$$

The parameter [S] in this equation denotes the acceptor or donor substrates. The kinetic parameters V_{max} and K_m were derived from the best fit of the Michaelis-Menten equation by using nonlinear regression with the GraphPad PRISM 3.0 program.

2.2.10 Crystallography

All mutants related to O^2 enzyme were crystallized using conditions similar to the native GTA and GTB enzymes (5). Data were collected on a Rigaku R-AXIS4++ at a distance of 100 mm and exposure times between 4.0–5.0 min for 0.5 degree oscillations. X-rays were produced on a MicroMax 30 W generator coupled to Osmic "Blue" confocal x-ray mirrors. The crystals were frozen at the cold head and maintained under cryogenic conditions at a temperature of -160 °C. Data sets of mutants for unliganded and in complexes with UDP-GalNAc and acceptor were solved using native GTA or GTB (Protein Data Bank accession code 1LZ0, 1LZ7, respectively) as a starting model with refinement using the program CNS (19) and modeling with the program SETOR (20).

2.3 Results

2.3.1 Site-directed mutagenesis

2.3.1.1 P74S mutant

The P74S mutant was prepared by PCR using the BAAA clone as a template. One round of PCR was sufficient to make the P74S mutant because the entire sequence could be covered by the forward primer (HJL01) and the reverse primer (PCR3B). The bright band showed that the size of product was correct (approximately 900 bp) (Fig. 2–5). The purified PCR product and pCWΔlac were digested with *Eco*RI and *Hin*dIII, then combined for a ligation reaction. Plasmids purified from individual *E. coli* transformants were digested with *Eco*RI and *Hin*dIII (Fig. 2–6). The size of digested product was correct. The lower band of Fig. 2–6 revealed that digested P74S product is approximately 900 bp. DNA sequencing analysis was performed to verify the mutation and the rest of the sequence of the mutant.

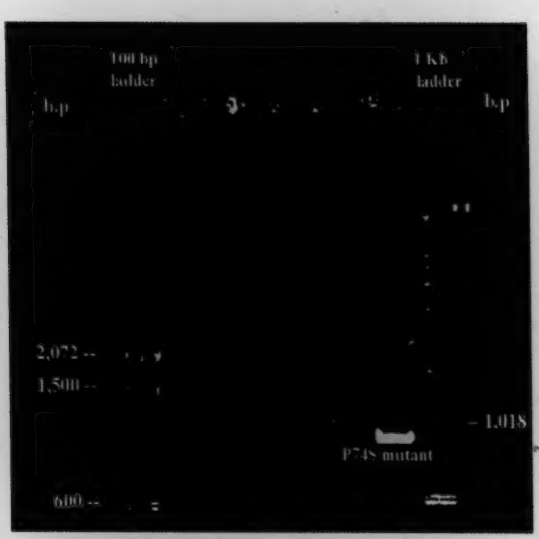

FIG. 2–5. **Agarose gel of PCR reaction for the P74S mutation.**

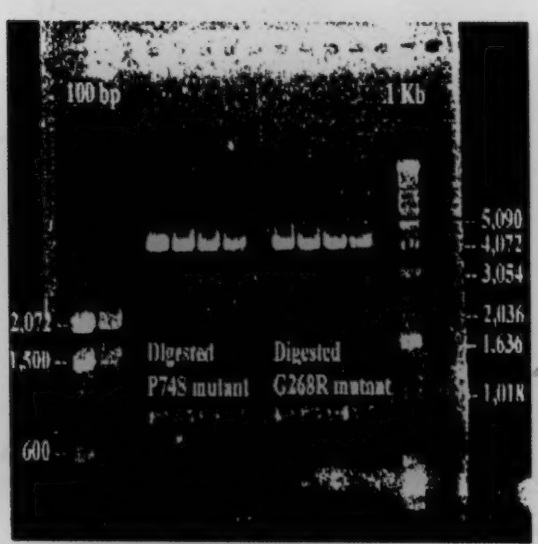

FIG. 2–6. **Agarose gel of digested mutants.** Upper and lower bands digested with *Eco*RI and *Hin*dIII represent the pCWΔlac plasmid and mutants, respectively.

2.3.1.2 G268R mutant

The G268R mutant was prepared by two rounds of PCR using the BAAA clone as a template. The first PCR was necessary to make a short (aa 53–271) and a long product (aa 266–354) and the subsequent PCR was required to make the full gene. The gel of each product of the first PCR revealed that the size of both the short and long products were correct (270 and 620 bp, respectively) (Fig. 2–7). In the second PCR, the short and long products were combined giving an approximately 900 bp gene product (Fig. 2–8). The purified PCR product and pCWΔlac were digested with *Eco*RI and *Hin*dIII then combined for the ligation reaction (Fig. 2–6). The lower band of Fig. 2–6 confirmed the right size of digested G268R product from the purified plasmid (approximately 900 bp).

2.3.1.3 P74S & G268R mutant

The P74S and G268R double mutant was prepared by PCR using the G268R mutant as a template. One round of PCR was sufficient to make the double mutant

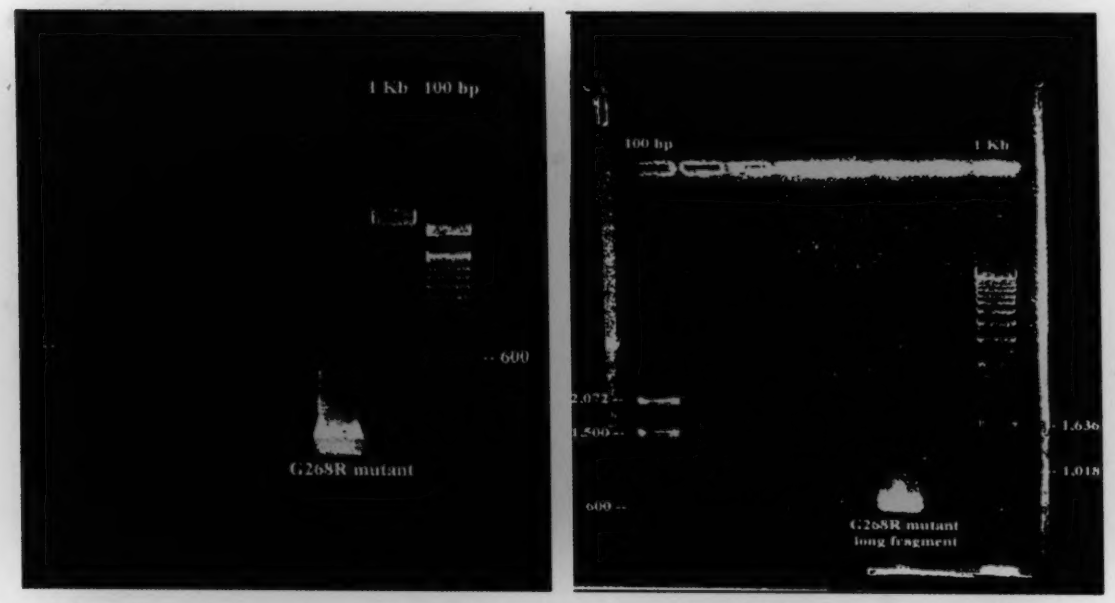

FIG. 2–7. **Agarose gel of the first PCR fragments for the G268R product.** The left gel shows the short fragment, the right gel the long fragment.

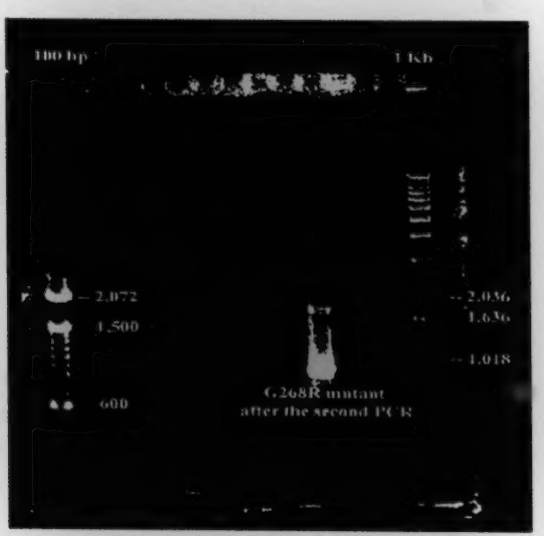

FIG. 2–8. **Agarose gel of the second PCR reaction for the G268R product.**

completely because the whole sequence could be covered by the forward primer (HJL01) and the reverse primer (PCR3B). A single band in agarose gel showed that the size of product was correct (approximately 900 bp) (Fig. 2–9). The purified PCR product and pCWΔlac were digested with *Eco*RI and *Hin*dIII then combined for the ligation reaction (Fig. 2–10). Plasmid purified from individual transformants were digested *Eco*RI and *Hin*dIII. The lower band of Fig. 2–10 confirmed the correct size of digested G268R product (approximately 900 bp).

2.3.2 Expression and purification

The soluble form of the mutants were expressed in *E. coli* and purified by a chromatography on SP-Sepharose Fast Flow ion-exchange column and a UDP-Hex affinity column. The purified enzymes, crude extracts, and all of the flow-through from the columns were monitored by SDS-PAGE (Fig. 2–11, 2–12, 2–13). The lanes show the crude lysate after French Press disruption (Lys), the supernatant from ultracentrifugation (S/N), the eluant from the SP-Sepharose FF column from loading the enzyme (SP-LFT), the eluant from the SP-Sepharose FF column when washing the column with

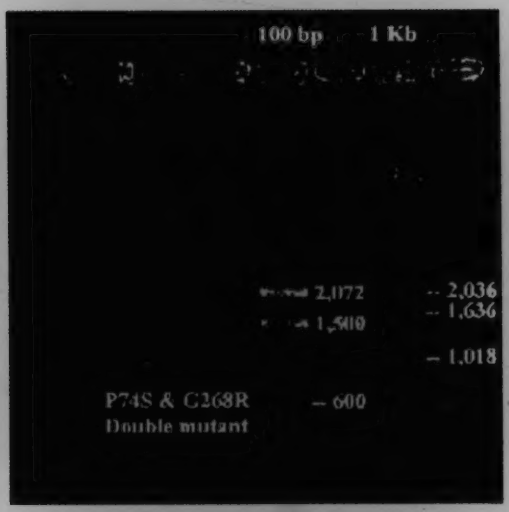

FIG. 2–9. **Agarose gel of the first PCR reaction of P74S & G268R product.**

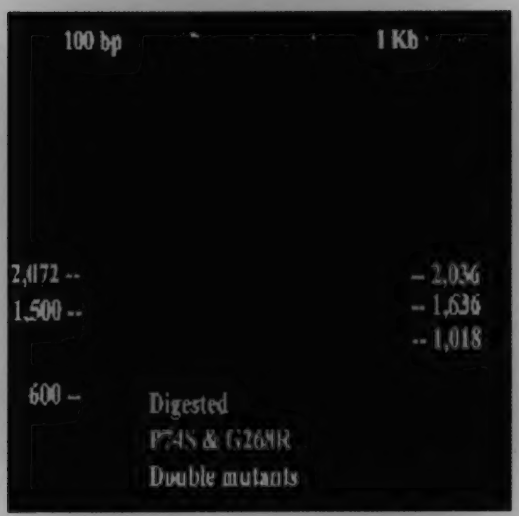

FIG. 2–10. **Agarose gel of digested P74S & G268R product.** Upper and lower bands represent the digested pCWΔlac and double mutant, respectively.

FIG. 2–11. **SDS-PAGE result for the P74S mutant.**

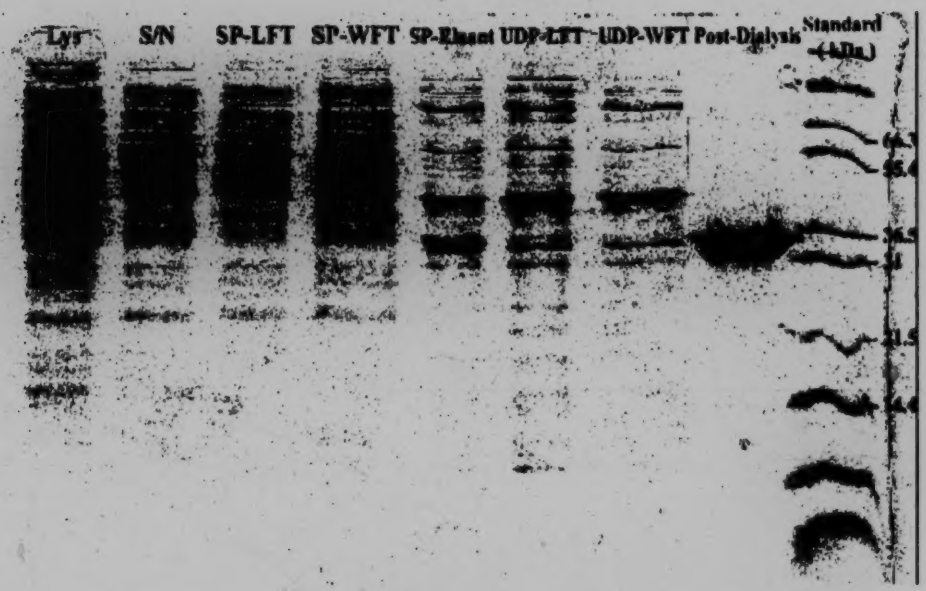

FIG. 2–12. **SDS-PAGE result for the G268R mutant.**

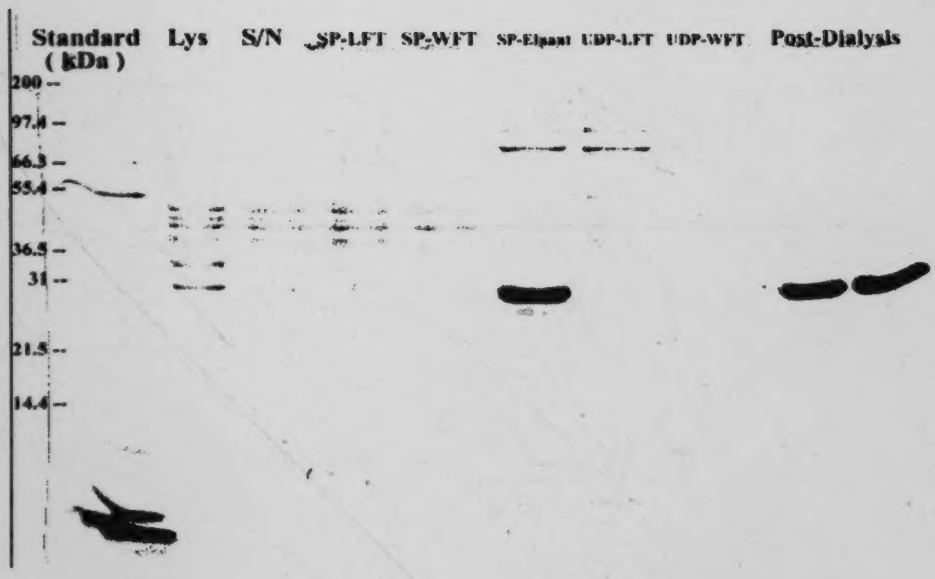

FIG. 2–13. **SDS-PAGE result for the P74S & G268R mutant (double mutation).**

equilibration buffer (SP-WFT), the eluant from the SP-Sepharose FF column when washed with equilibration buffer containing 0.5 M NaCl (SP-eluant), the eluant from the UDP-Hex column when loading the enzyme (UDP-LFT), the eluant from the UDP-Hex column when washing the column with equilibration buffer (UDP-WFT) and the enzyme after desorption from the UDP-Hex column and dialysis to remove UDP (Post-dialysis). Molecular mass standards are also shown on the gels. The final enzymatic activity was estimated with the various saturating GTA assays (Table 2–1). The P74S mutant had a specific activity of 22.2 Unit/mg, the G268R mutant and the P74S & G268R double mutants showed 1.3×10^{-3} and 1.0×10^{-4} Unit/mg, respectively using the saturating GTA assay where a unit is the amount of enzyme that converts 1 μmol of substrate to product per minute at 37 °C. Expression levels for all of the mutants were good and yields of the final purified proteins ranged from 26.3 mg/liter for G268R mutant to 107.9 mg/liter of *E. coli* for the P74S double mutant.

2.3.3 Kinetic characterization

The calculated molecular mass (in Da) of the mutants were compiled

TABLE 2–1.

Relative rates of transfer of the natural donor UDP-GalNAc by mutant glycosyltransferases

Critical aa residues and mutations					A-activity
176	235	266	268	Mutation	*Unit/mg*
A_{Arg}	A_{Gly}	A_{Leu}	A_{Gly}	N/A[a]	22.1
B_{Gly}	A_{Gly}	A_{Leu}	A_{Gly}	N/A	54.0
B_{Gly}	A_{Gly}	A_{Leu}	A_{Gly}	Pro74Ser	22.2
BAAA				Gly268Arg	1.3×10^{-3}
BAAA				Pro74Ser & Gly268Arg	1.0×10^{-4}

[a] N/A denotes no mutation

(Table 2–2). These molecular mass values were used for k_{cat} calculations. Kinetic constants were determined for each mutant using the A donor (UDP-GalNAc) and the B donor (UDP-Gal) with seven acceptor or donor concentrations at a saturating concentration of the alternate substrate. The k_{cat} for the BAAA P74S mutant was comparable to that of the wild-type GTA, but lower than that of the BAAA mutant (Table 2–3). The specificity constant (k_{cat}/K_m) of P74S mutant was decreased 9-fold for the acceptor and 16-fold for the A donor compared to wild-type GTA (Table 2–3). The BAAA G268R mutant showed very low A activity. The k_{cat} value was 125,000 times less than that of the BAAA and 33,000 times less than that of the BAAA P74S mutant. The specificity constant (k_{cat}/K_m) also showed a large decrease, 430,000-fold for the acceptor and 700,000-fold for the donor compared to that of the BAAA mutant, and 64,000-fold for the acceptor and 200,000-fold for the donor compared to BAAA P74S mutant. The BAAA P74S & G268R mutant showed the lowest A activity of all of the enzymes. The k_{cat} was only $3.4\times 10^{-5}\ s^{-1}$, and the k_{cat}/K_m was 8.1×10^{-5}, 1.5×10^{-5} for the acceptor and donor, respectively. These values are 1.6×10^{6}, 1.6×10^{6},and 3.0×10^{7}-fold smaller than those of the BAAA mutant for k_{cat}, k_{cat}/K_A, k_{cat}/K_B, respectively, and 4.3×10^{5}-fold smaller than the k_{cat} of the BAAA P74S mutant, and 2.4×10^{5}-fold, and 8.6×10^{6}-fold smaller than

TABLE 2–2.

Calculated molecular mass of the mutants

Mutants	Calculated Molecular Mass *Da*
-10/BAAA/- P74S	33,904.35
-10/BAAA/- G268R	34,013.48
-10/BAAA/- P74S & G268R	34,003.48

TABLE 2–3.

Kinetic constants of mutant glycosyltransferases using UDP-GalNAc

Critical residues and mutants	K_A[a]	K_B[b]	k_{cat}	k_{cat}/K_A	k_{cat}/K_B
176 235 266 268	μM	μM	s^{-1}	$mM^{-1}s^{-1}$	$mM^{-1}s^{-1}$
$A_{Arg}A_{Gly}A_{Leu}A_{Gly}$	9.9	8.7	17.5	1770	2010
BAAA[c]	43	126	55	1300	440
P74S[d]	75	113	14.6	194	129
G268R	157	700	0.00044	0.003	0.00063
P74S & G268R	42	2323	0.000034	0.00081	0.000015

[a] K_A is the Michaelis-Menten constant for the acceptor (Fucα1–2Galβ-O-$(CH_2)_7CH_3$).

[b] K_B is the Michaelis-Menten constant for the donor (UDP-GalNAc).

[c] Data for enzyme with 10 additional amino acids at the N-terminal.

[d] All mutations were based on the BAAA clone as a template.

the k_{cat}/K_m for the acceptor and donor, respectively. The kinetic constant for the double mutated enzyme (BAAA P74S & G268R) was also compared to that of the BAAA G268R single mutant. The k_{cat} of the double mutant showed a 13-fold decrease, and the k_{cat}/K_m values for the acceptor and donor showed a 3.7 and 20-fold decrease compared to the G268R mutant (Table 2–3). All the kinetic data are rearranged by ratio in Table 2–4. Kinetic studies with the B donor (UDP-Gal) were only carried out for the BAAA P74S mutant since transfer rates of the BAAA G268 mutant and the BAAA P74S & G268R were too low to measure (Table 2–5). The k_{cat} of the BAAA P74S mutant showed a 60-fold decrease compared to that of the wild-type GTB, but was 2-fold higher than that of the BAAA mutant. The k_{cat}/K_m value for the acceptor was similar to that of the BAAA, and showed a 3-fold decrease for the donor. All the kinetic data were rearranged by ratio normalized to wild-type GTB (Table 2–6).

2.3.4 Crystal structure

The crystal structure of the mutants related to the O^2 glycosyltransferases provides structural features for reduced enzymatic activity. Table 2–7 shows the statistical

TABLE 2–4.

Kinetic constants ratio for mutant glycosyltransferases using UDP-GalNAc

Critical residues and mutants	K_A[a]	K_B[b]	k_{cat}	k_{cat}/K_A	k_{cat}/K_B
176 235 266 268	μM	μM	s^{-1}	$mM^{-1}s^{-1}$	$mM^{-1}s^{-1}$
$A_{Arg}A_{Gly}A_{Leu}A_{Gly}$	1.0[d]	1.0	1.0	1.0	1.0
BAAA[c]	4.34	14.48	3.14	0.73	0.22
P74S[d]	7.57	12.98	0.83	0.11	0.064
G268R	15.85	80.45	2.5×10^{-5}	1.7×10^{-6}	8.1×10^{-7}
P74S & G268R	4.24	267	0.2×10^{-5}	4.6×10^{-7}	0.7×10^{-8}

[a] K_A is the Michaelis-Menten constant for the acceptor (Fucα1–2Galβ-O-$(CH_2)_7CH_3$).

[b] K_B is the Michaelis-Menten constant for the donor (UDP-GalNAc).

[c] All mutations were based on the BAAA as a template.

[d] Kinetics constants for wild-type GTA was set as a one for comparison.

[e] Data for enzyme with 10 additional amino acids at the N-terminal.

TABLE 2–5.

Kinetic constants of mutant glycosyltransferases using UDP-Gal

Critical residues and mutants	K_A[a]	K_B[b]	k_{cat}	k_{cat}/K_A	k_{cat}/K_B
176 235 266 268	μM	μM	s^{-1}	$mM^{-1}s^{-1}$	$mM^{-1}s^{-1}$
B_{Gly} B_{Ser} B_{Met} B_{Ala}	88	27	5.1	58	190
BAAA[c]	68	43	0.037	0.54	0.86
P74S[d]	141	278	0.08	0.57	0.29
G268R	N/A[e]	N/A	N/A	N/A	N/A
P74S & G268R	N/A	N/A	N/A	N/A	N/A

[a] K_A is the Michaelis-Menten constant for the acceptor (Fucα1–2Galβ-O-$(CH_2)_7CH_3$).

[b] K_B is the Michaelis-Menten constant for the donor (UDP-Gal).

[c] N/A, values are too low to measure.

[d] All mutations were based on the BAAA as a template.

[e] Data for enzyme with 10 additional amino acids at the N-terminal.

TABLE 2–6.

Kinetic constants ratio for mutant glycosyltransferases using UDP-Gal

Critical residues and mutants	K_A[a]	K_B[b]	k_{cat}	k_{cat}/K_A	k_{cat}/K_B
176 235 266 268	μM	μM	s^{-1}	$mM^{-1}s^{-1}$	$mM^{-1}s^{-1}$
B_{Gly} B_{Ser} B_{Met} B_{Ala}	1	1	1	1	1
BAAA[d]	0.77	1.59	0.007	0.0093	0.0045
P74S	1.60	10.30	0.016	0.0098	0.0015
G268R	N/A[c]	N/A	N/A	N/A	N/A
P74S & G268R	N/A	N/A	N/A	N/A	N/A

[a] K_A is the Michaelis-Menten constant for the acceptor (Fucα1–2Galβ-O-$(CH_2)_7CH_3$).

[b] K_B is the Michaelis-Menten constant for the donor (UDP-Gal).

[c] N/A, values are too low to measure.

[d] All mutations were based on the BAAA as a template.

TABLE 2–7.

Data collection and refinement results for O^2 related mutants.

	P74S	G268R	P74S&G268R
Resolution (Å)	19.87–1.59	27.38–1.59	19.94–1.49
Space group	$C222_1$	$C222_1$	$C222_1$
a (Å)	52.7	52.7	52.9
b (Å)	150.1	150.1	150.3
c (Å)	79.7	79.7	79.8
R_{merge} (%)[a,b]	4.9 (39.0)	3.4 (30.2)	3.7 (29.3)
Completeness (%)[a]	99.3 (98.6)	93.5 (78.7)	91.6 (54.5)
Unique reflections	42,625	40,458	47,938
Refinement			
Resolution	20–1.8	20–1.8	20–1.8
R_{work} (%)[c]	20.9	20.9	20.7
R_{free} (%)[d]	24.4	24.3	23.8
r.m.s.[e] bond (Å)	0.006	0.005	0.005
r.m.s. angle (°)	1.30	1.30	1.32

[a] Values in parentheses represent high resolution shell.

[b] $R_{merge} = \Sigma|I_{obs} - I_{ave}|/\Sigma I_{ave}$

[c] $R_{work} = \Sigma||F_o| - |F_c||/\Sigma|F_o|$

[d] 10% of reflections were omitted in R_{free} calculations.

[e] r.m.s., root mean square.

results of three mutants grown unliganded and Table 2–8 shows the crystallographic result of the same mutants in complex with donor and acceptor substrates. All structures showed excellent electron densities with the exception of the disordered loop (aa 176–195) and the final 10 amino acid residues at the C-terminal, which were also absent in the native glycosyltransferase structures. The electron density map around the active site of the O^2 enzyme was satisfactory (Fig. 2–14).

2.4 Discussion

In this chapter, human blood group O^2 enzyme was produced and characterized. It is a triple mutant of GTA with arginine 176 replaced with glycine, proline 74 replaced with serine, and glycine 268 replaced with arginine. Except for the BAAA P74S mutant, the purified mutants for human glycosyltransferase O^2 from *E. coli* showed very low enzymatic activity. The O^2 glycosyltransferase shares some features with GTA because it is a soluble form and the sequence is identical except for three amino acid residues at positions 74, 176, and 268. However, the O^2 glycosyltransferase is an inactive enzyme. To determine which residue has a dominant effect on the transferase activity, three

TABLE 2–8.

Data collection and refinement results for O^2 related mutant crystals grown in the presence of donor substrates and acceptors

	P74S+DA[a]+ UDP-GalNAc	G268R+HA+ UDP-GalNAc	P74S&G268R+HA[b]+ UDP-Gal
Resolution (Å)	19.74–1.49	19.96–1.55	19.88–1.55
Space group	$C222_1$	$C222_1$	$C222_1$
a (Å)	52.6	52.8	52.7
b (Å)	149.1	150.1	150.3
c (Å)	79.1	79.8	79.5
R_{merge} (%)[c,d]	6.2 (27.2)	5.3 (36.5)	3.2 (25.9)
Completeness (%)[c]	95.0 (61.0)	95.0 (72.1)	96.6 (74.1)
Unique reflections	48,630	44,096	44,636
Refinement			
Resolution	20–1.9	20–1.8	20–1.8
R_{work} (%)[e]	21.0	20.7	20.3
R_{free} (%)[f]	25.3	23.4	23.2
r.m.s.[g] bond (Å)	0.006	0.005	0.005
r.m.s. angle (°)	1.32	1.33	1.31

[a] DA, Deoxy-acceptor

[b] HA, H-acceptor.

[c] Values in parentheses represent high resolution shell.

[d] $R_{merge} = \Sigma|I_{obs} - I_{ave}|/\Sigma I_{ave}$

[e] $R_{work} = \Sigma||F_o| - |F_c||/\Sigma|F_o|$

[f] 10% of reflections were omitted in R_{free} calculations.

[g] r.m.s., root mean square.

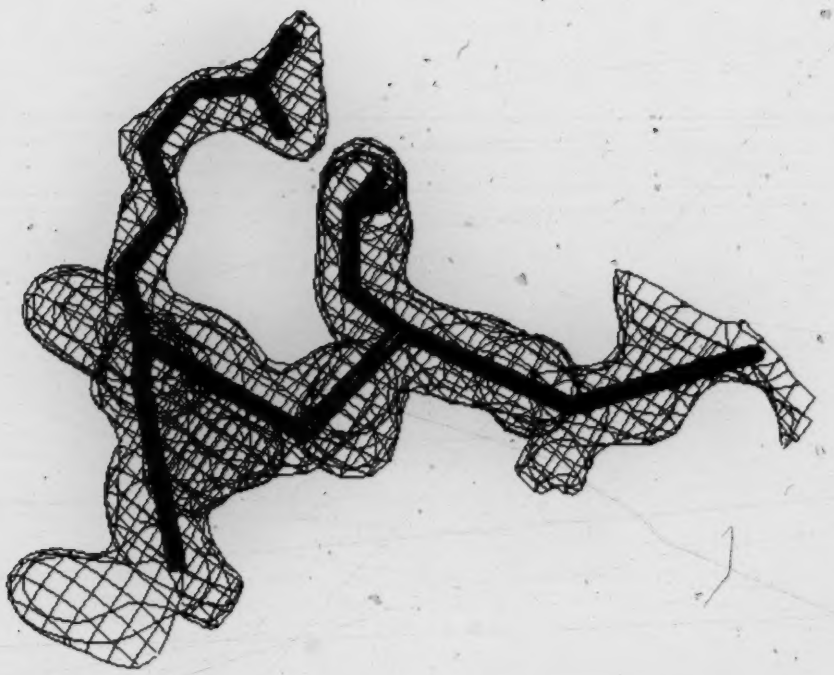

FIG. 2–14. **The electron density map of the active site of the O^2 enzyme.** Substitution of G268R blocks the active site completely. *Red* represents Arg-268. Leu-266 is in *orange* and limited main chains (264–270) are shown. The electron density map was generated using SETOR (20).

mutants (P74S, G268R and double) were prepared, purified and analyzed by enzyme kinetics and crystallography. Except for the P74S mutant related to the O^2 glycosyltransferase, the mutants were essentially non-functional.

The BAAA P74S mutant showed a similar A activity to that of the wild type GTA, but lower A activity compared to that of the BAAA mutant. The K_B values of these two mutants were similar, and the K_A values had small differences. The k_{cat}, however, showed a 4-fold decrease compared to that of the BAAA mutant. The low enzymatic catalytic efficiency based on the k_{cat}/K_m data showed that P74S mutant had a lower frequency for enzyme and substrate encounters than the frequency of the BAAA mutant, and lower frequency than that of the wild-type GTA. The mutation at residue 74 in BAAA appeared to have a relatively small effect on the donor binding, with more of an effect on the acceptor and it significantly affected the enzyme turnover (Table 2–4). The proline residue at the codon 74 is located in the folded peptide region of the glycosyltransferase. Presumably, the mutation of proline 74 to serine changed the topology of a β-sheet at the N-terminal. As a result, the topology of the enzyme was changed to induce unfavorable interaction with a substrate compared to the interaction of a BAAA enzyme without this mutation.

The BAAA G268R mutant showed very low activity with a high K_m for the acceptor and donor. The K_A value was 3.7-fold higher than that of the BAAA mutant, and the K_B value was 5.5 times higher than that of the BAAA. These results indicate that the mutation at residue 268 has some affect on the donor and acceptor binding. The k_{cat}, however, was markedly decreased. The value showed a 40,000-fold lower k_{cat} compared with that of the wild-type GTA. The k_{cat}/K_m data showed a 600,000-fold and 3,000,000-fold decrease of catalytic efficiency compared to that of the wild-type GTA for the acceptor and donor, respectively. The arginine residue is much bulkier than a Gly residue. As a result of this mutation, the crucial sugar-nucleotide binding site was altered with no space for the monosaccharide of donor giving a virtually inactive enzyme.

The BAAA P74S & G268R double mutant enzyme showed strikingly low enzymatic activity. The k_{cat} was 500,000-fold smaller than that of the wild-type GTA, and the K_m for donor was extremely high compared with that of the wild-type GTA (4-fold and 267-fold higher K_A and K_B, respectively). This k_{cat} was 13 times smaller value than that of the G268R mutant. This result shows that the mutations at both residue Pro-74 and Arg-268 contributed to non-functionality of the O^2 enzyme. Both the unfavorable topology due to the mutation of proline to serine 74 and the disruption of the important

sugar-nucleotide binding site by steric hindrance at codon 268 were responsible for the loss of transferase activity.

Crystallographic analysis revealed that Arg-268 has a dominant effect on the enzymatic activity of O^2 glycosyltransferase. The large arginine side chain at the active site blocks access of the donor substrate completely (Fig. 2–14). The structure of O^2 enzyme with donor and acceptor substrates shows the effect of Arg-268 (Fig. 2–15).

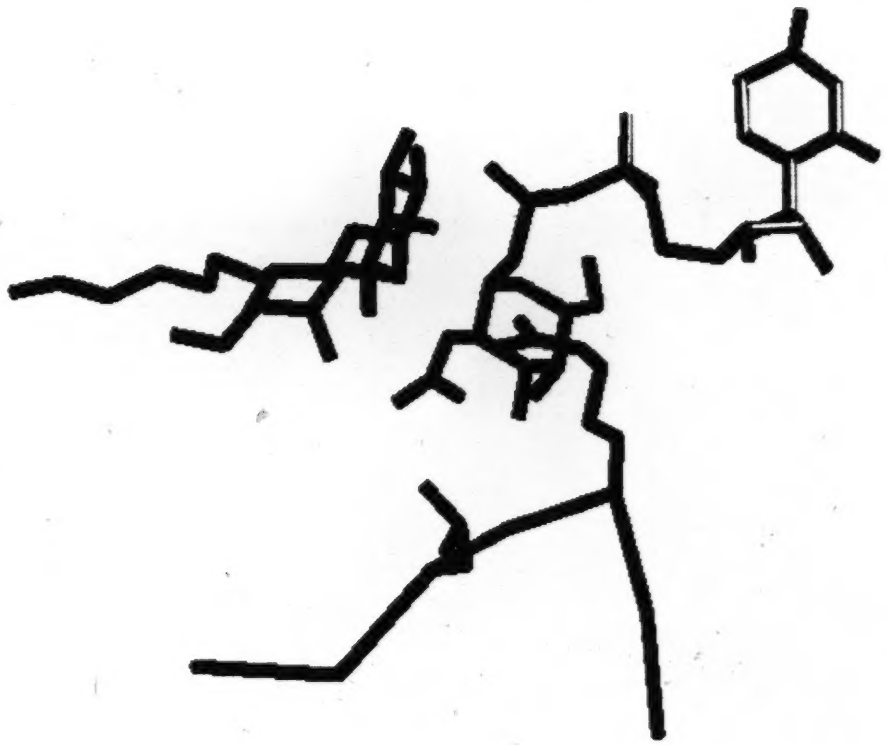

FIG. 2–15. **The structure of the active site of the O^2 enzyme with donor and acceptor substrates.** The altered active site cleft of the O^2 enzyme precluded substrate binding due to the large size of arginine 268 (*Red*). *Orange* represents Leu-266. Limited main chains (264–270) are shown. UDP-GalNAc is represented by *yellow* and *blue* represents the H-acceptor. The GalNAc moieties were modeled by SETOR (20).

Acknowledgements

Assistance with DNA sequencing was provided by the Molecular Biology Service Unit, Department of Biological Sciences, University of Alberta. We wish to thank F. W. Dahlquist for the pCW vector, W. Wakarchuk for the pCWΔlac vector, and O. Hindsgaul for the Fucα1–2Galβ-O-$(CH_2)_7CH_3$ acceptor. Funding for this work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC), Alberta Ingenuity Centre for Carbohydrate Science (AICCS) and the Canadian Institute of Health Research (CIHR).

References

1. Watkins, W. M. (1980) *Adv. Hum. Genet.* **10,** 379–385

2. Palcic, M. M., Seto, N. O., and Hindsgaul, O. (2001) *Transfus. Med.* **11,** 315–323

3. Yamamoto, F., Clausen, H., White, T., Marken, J., and Hakomori, S. (1990) *Nature* **345,** 229–233.

4. Busch, C., Hofmann F., Selzer, J. Munro, S., Jeckel, D., and Aktories, K. (1998) *J. Biol. Chem.* **273,** 19566–19572

5. Patenaude, S. I., Seto, N. O. L., Borisova, S. N., Szpacenko, A., Marcus, S. L., Palcic, M. M., and Evans, S. V. (2002) *Nat. Struct. Biol.* **9,** 685–690

6. Olsson, M. L., and Chester, M. A. (2001) *Transfus. Med.* **11,** 295–313

7. Yamamoto, F., McNeill, P. D., Yamamoto, M., Hakimori, S., Bromilow, I. M., and Duguid, J. K. (1993) *Vox Sang.* **64,**175–178

8. Yamamoto, F., and McNeill, P. D. (1996) *J. Biol. Chem.* **271,** 10515–10520

9. Amando, M., Bennett, E. P., Carneiro, F., and Clausen, H. (2000) *Vox Sang.* **79,** 219–226

10. Grunnet, N., Steffensen R., Bennett, E. P., and Clausen, H. (1994) *Vox Sang.* **67,** 210–215

11. Amado, M., Bennett, E. P., Carneiro, F., and Clausen, H. (2000) *Vox Sang.* **79,** 219–226

12. Greenwell, P., and Watkins, M. M. (1987) *Proc. IXth Int. Symp. on Glycoconjugates*, Montreuil, pE34

13. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (eds) (1997) Current Protocols in Molecular Biology, Vol. 1, pp.8.5.7–10, John Wiley and Sons, Inc., New York

14. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, New York

15. Seto, N. O. L., Palcic, M. M., Hindsgaul, O., Bundle, D. R., and Narang, S. (1995) *Eur. J. Biochem.* **234,** 323–328

16. Seto, N. O. L., Compston, C. A., Evans, S. V., Bundle, D. R., Narang, S. A., and Palcic, M. M. (1999) *Eur. J. Biochem.* **259,** 770–775

17. Gegner, J. A., and Dahlquist, F. W. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 750–754

18. Bradford, M. M. (1976) *Anal. Biochem.* **72,** 248–254

19. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sec. D.* **54,** Pt. 5, 905–921

20. Evans, S. V. (1993*) J. Mol. Graphics* **11,** 134–138

Chapter 3. Characterization of Recombinant *cis*-AB Glycosyltransferase

3.1 Introduction

The ABO blood group glycosyltransferases utilize the donor substrates UDP-GalNAc or UDP-Gal, respectively, in the biosynthesis of the A and B antigens (1, 2). The A synthesizing α1–3 *N*-acetylgalactosaminyltransferase (GTA, EC 2.4.1.40) transfers GalNAc from UDP-GalNAc to Fucα1–2Galβ-R terminated blood group H acceptors, producing the A antigen GalNAcα1–3[Fucα1–2]Galβ-R. Similarly, the B synthesizing α1–3 galactosyltransferase (GTB, EC 2.4.1.37) transfers Gal from UDP-Gal to the same acceptor, producing the B antigen Galα1–3[Fucα1–2]Galβ-R (Fig. 3–1). These reactions are regio- and stereospecific and occur with retention of configuration at the anomeric center of the transferred monosaccharide.

GTA and GTB are highly homologous enzymes differing in only four of 354 residues (Fig. 3–2), (3). A change of these four critical amino acids (Arg-176 → Gly, Gly-235 → Ser, Leu-266 → Met, and Gly-268 → Ala) converts the specificity from GTA to GTB. X-ray diffraction studies of crystals of GTA and GTB with their substrate

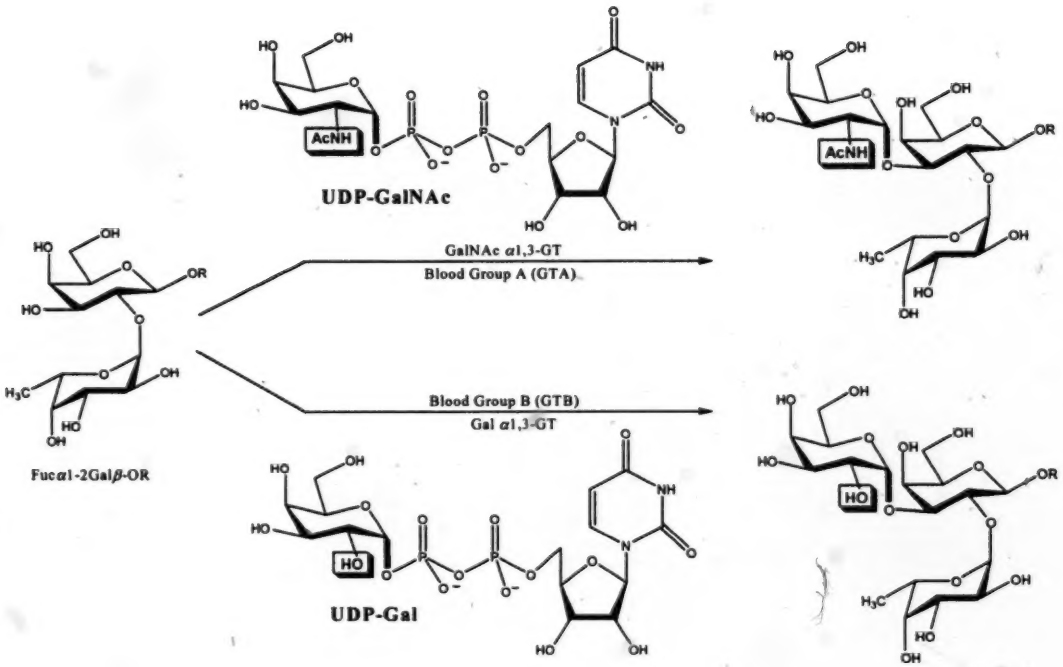

FIG. 3–1. **The reactions catalyzed by the human blood group A and B glycosyltransferases.**

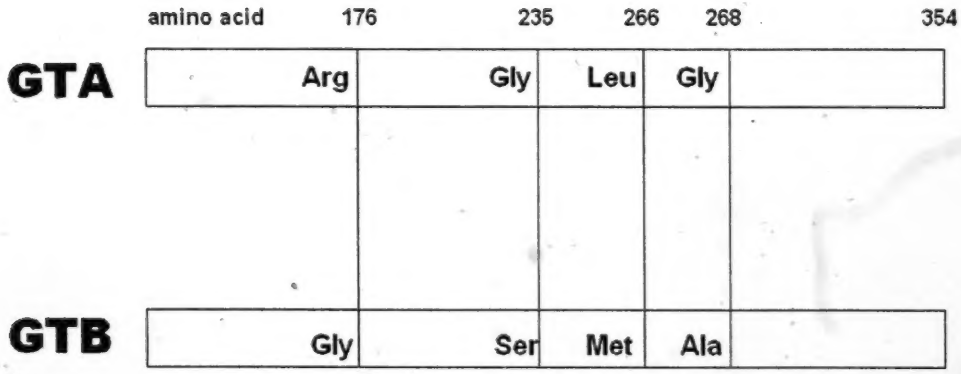

FIG. 3–2. **Schematic representation of the divergent amino acid positions for GTA and GTB.** Critical amino acid residues are shown. Alteration of these critical residues determines the specificity of donor utilized.

complexes have shown the structural basis for substrate binding and residues involved in donor selection (4).

Since Yamaguchi reported his discovery of a *cis*-AB phenotype that had weak enzyme activity and expressed both A and B antigens (5), numerous genetic studies on *cis*-AB have been carried out. These are dual specificity enzymes that utilize both UDP-Gal and UDP-GalNAc as donors. Yamamoto analyzed the molecular genetic basis of the *cis*-AB allele and showed it was an intermediate form of the ABO gene defined as AAAB, where this four-letter code is based on whether the critical amino acid in GTA or GTB. In this code GTA is designated AAAA (Arg-176, Gly-235, Leu-266, Gly-268), GTB is BBBB (Gly-176, Ser-235, Met-266, Ala-268) and AAAB is Arg-176, Gly-235, Leu-266, Ala-268 (6). This result was confirmed by Fukumori and Ogassawara (7, 8). Mifsud found a different type of *cis*-AB that also showed a dual functionality (BBAB) (9). Later *cis*-AB enzymes, having a BABB, ABAB and BAAB constellation at the four critical amino acid positions were discovered (6, 9, 10). A different type of the *cis*-AB enzyme was reported that was a full-length form of GTB with proline 234 replaced with alanine (11). More recently, a *cis*-AB enzyme was discovered that was a full-length form of GTA with Leu-266 replaced by a Gly and Gly-268 replaced by Ala. This type of *cis*-AB

enzyme prevails in the mouse population and encodes for an enzyme with both A and B transferase activities (12). All of the *cis*-AB enzymes described above are shown in Fig. 3–3.

While previous studies have shown that *cis*-AB enzymes have dual enzymatic activity utilizing both A and B donors (6–12), no kinetic characterizations have not yet been carried out. In this chapter a recombinant *cis*-AB glycosyltransferase was prepared using site-directed mutagenesis, and purified for kinetic characterization to confirm its dual functionality. Preparative scale syntheses of the A and B antigens were carried out with the enzyme. Finally, X-ray crystallography was used to determine its structure.

3.2 Experimental procedures

3.2.1 Materials and general techniques

All molecular biology procedures were carried out according to standard procedures (13, 14). Tryptone (Bacto); yeast extract and casamino acids (Difco); glycerol

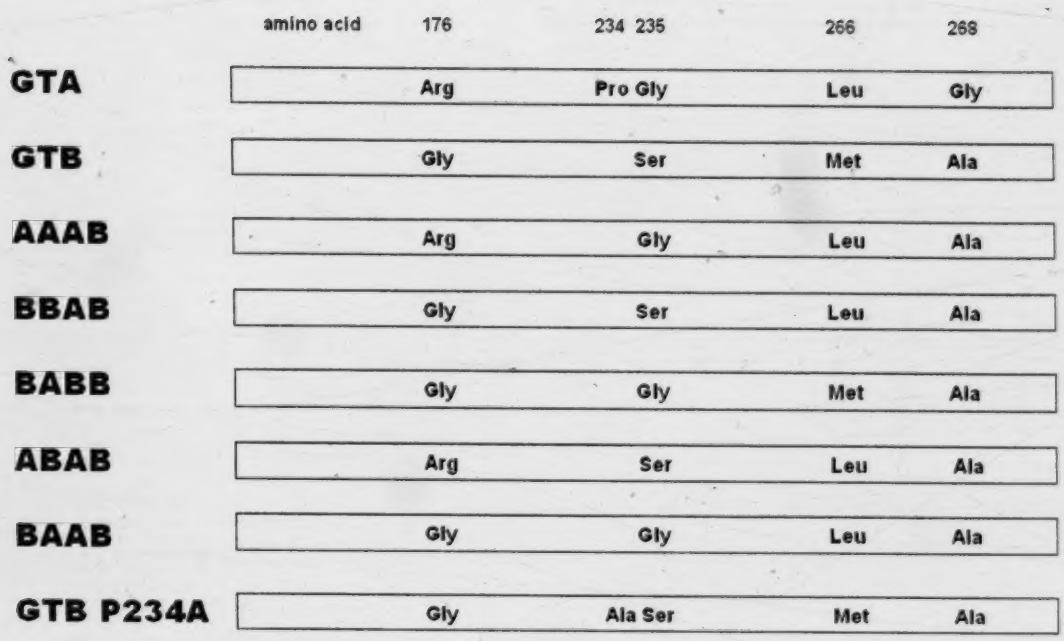

FIG. 3–3. **Schematic diagram of the changes in amino acids that give *cis*-AB enzymes.**

(GibcoBRL); Na_2HPO_4, dithiothreitol (DTT), 3-(*N*-morpholino) propanesulfonic acid (MOPS) (Fisher); isopropyl-1-thio-β-D-galactopyranoside (IPTG) (Rose Scientific); protease inhibitor cocktail tablets (Complete, EDTA-free) (Roche); PCR reagents including *Pfx* DNA polymerase, *Pfx* buffer, and DNA ladder (Invitrogen); restriction enzymes and DNA-modifying enzymes (New England Biolabs and Life Technologies, Inc.); primers including HJL01, HJL06, HJL07 (IDT); KH_2PO_4, NH_4Cl, NaCl, $MnCl_2$, $CaCl_2$, thiamine hydrochloride (vitamin B1), D[-]-α-aminobenzylpenicillin (ampicillin), ethylenediaminetetraacetic acid (EDTA-tetrasodium salt: hydrate), glucose, UDP, UDP-GalNAc, and UDP-Gal, GenElute Plasmid Miniprep Kit, Post-Reaction Purification Columns (Sigma); QIAquick PCR Purificatioin Kit, QIAquick Gel Extraction Kit (Qiagen); BL21-gold competent cells (Stratagene); DYEnamic ET terminator cycle sequencing kit (Amersham Biosciences); Mini-sub Gel GT for Agarose gel electrophoresis for PCR reaction and protein assay dye reagent concentrate (Bio-Rad); the media for SP-Sepharose Fast Flow and UDP-hexanolamine columns (Pharmacia Fine Chemicals); the radioactive [6-^{3}H]-labeled analogs of UDP-GalNAc and UDP-Gal (American Radiolabeled Chemicals); Sep-Pak C_{18} (+) reverse phase cartridges (Waters); Ecolite (+) liquid scintillation cocktail (ICN); Centriplus-10 protein concentrators

(Amicon); bovine gamma globulin (IgG), Slide-a-Lyzer dialysis cassettes (Pierce Chemical Co.); SDS-PAGE system (Hoefer Scientific Instruments); thin layer chromatography (TLC) plates (E. Merch AG) were purchased commercially. The acceptor substrate Fucα1–2Galβ-O-$(CH_2)_7CH_3$ was a gift from Dr O. Hindsgaul (University of Alberta). A Rigaku R-AXIS4++ detector, MicroMax X-ray generation system (Rigaku/MSC) and Osmic "Blue" confocal x-ray mirrors (Osmic) were used for the crystallography.

3.2.2 Cloning of L266G

The original GTB and GTA gene sequences (aa 54–354) were described in previous papers (15, 16). In this chapter, GTB and GTA (aa 63–354) are denoted as wild-type enzymes, and the AAGlyB mutant (Arg-176, Gly-235, Gly-266, Gly-268) was constructed by PCR using the AAAB clone as a template. The forward primer MIN2 (5'-ATA TGA ATT CAT GGT TTC CCT GCC GCG TAT GGT TTA CCC GCA GCC GAA-3') introduced an *Eco*RI site at the 5' end, and the reverse primer PCR3B (5'- ATA ATT AAG CTT CTA TCA CGG GTT ACG AAC AGC CTG GTG GTT TTT- 3') introduced a

*Hin*dIII site at the 3' end of the gene. Two fragments were amplified with *Pfx* DNA polymerase (Invitrogen) by using the forward primer MIN2 together with HJL06 (5'- GAA AGC ACC ACC GTA GTA GAA GTC ACC TTC G- 3') and the reverse primer PCR3B with HJL07 (5'- C TAC TAC GGT GGT GCT TTC TTC GGT GGT TCC- 3'). HJL06 and HJL07 were designed so that the two fragments overlapped each other and have a single codon substitution (CTG to GGT) at codon 266. The two overlapping fragments were isolated, annealed by 3' extension by using PCR and amplified by using the outside primers MIN2 and PCR3B. The amplified genes were digested by restriction enzymes (*Eco*RI, *Hin*dIII) and ligated into the previously digested pCWΔlac vector (17). The ligation reaction was incubated at room temperature overnight and transformed into *E. coli* BL21-gold using $CaCl_2$ competent cells. A single transformant was inoculated into LB broth containing ampicillin and incubated overnight at 37 °C. Plasmids were purified with a mini plasmid preparation column. The entire sequence was confirmed by sequencing using a DYEnamic ET terminator cycle sequencing kit.

3.2.3 Purification of L266G

E. coli cells were grown overnight, at 30 °C, with shaking at 200 rpm, in 500 ml baffled flasks, in a starter culture containing 1 × TB broth (13.2 g of tryptone, 26.4 g of yeast extract, and 4.4 ml of glycerol/L), 1 × M9 salts (6.6 g of Na_2HPO_4, 3.3 g of KH_2PO_4, 1.1 g of NH_4Cl, and 0.55 g of NaCl/L), and M9 supplements (1 mM $MgCl_2$, 0.1 mM $CaCl_2$, 5 μg/ml vitamin B1, 0.2 % glucose, 0.4 % casamino acids, and 0.1 mg/ml ampicillin/L). The cultures were scaled up to 4 × 1 liter size (each 1 liter culture was in a 4 liter baffled flask) by using 30–100 ml of starter culture per 1 liter of scale-up broth (with the same concentration of all the materials as that of the starter broth), and further shaken at 200 rpm, 30 °C. The cultures were induced with 1 mM IPTG at OD_{600} 0.6–1.0 and harvested 20–24 h later. The cells were then spun down at 11,325 × *g* for 20 min, at 4 °C, to remove the supernatant. The pelleted cells were then frozen at -20 °C for purification later or resuspended immediately in 50 mM MOPS, 1 mM DTT, pH 7.0, on ice. A complete EDTA-free protease inhibitor cocktail tablet was added to the cell suspension, which was then passed once through a French Press at 20,000 psi and centrifuged at 105,000 × *g* for 60 min at 4 °C. The resulting supernatant was then loaded

onto an SP-Sepharose FF column (220 ml, 5.8 cm × 8.3 cm) at a flow rate of 5 ml/min. The SP-Sepharose FF column was regenerated between enzyme purifications by washing with 8 M urea, 2 M NaCl, 1 M NaOH and 30% (v/v) of isopropanol. The packing resin was assayed for activity to ensure that no contaminating enzymes were present. The column was equilibrated with 50 mM MOPS, 1 mM DTT, pH 7.0, prior to loading of the supernatant from ultracentrifugation. After loading, the column was washed with 50 mM MOPS, 1 mM DTT, pH 7.0, at 5 ml/min until OD_{280} was at the baseline (approximately 3 column volumes). The enzyme was eluted with 50 mM MOPS, 1 mM DTT, 0.5 M NaCl, pH 7.0, at 4 ml/min collecting 12 ml fractions. The fractions with appreciable amounts of enzymatic activity, as assessed by a standard assay for GTA activity (described below), were then pooled for loading onto a UDP-Hex column. Manganese chloride was added to the pooled fractions to a final concentration of 5 mM prior to loading onto the UDP-Hex column (20 ml, 1.8 cm × 8.2 cm). The UDP-Hex column was also regenerated by washing with 8 M urea between consecutive enzyme purifications to avoid contamination, and was equilibrated with 50 mM MOPS, 1 mM DTT, 0.5 M NaCl, 5 mM $MnCl_2$, pH 7.0, at a flow rate of 1–2 ml/min, 4 °C. Loading was done at a flow rate of 0.2–0.4 ml/min, overnight, followed by a wash with the equilibration buffer at a flow rate of 0.6–0.8

ml/min until OD_{280} was at the baseline. The enzyme was eluted at a flow rate of 1 ml/min collecting 4 ml fractions with the loading buffer containing 10 mM UDP. The fractions were assayed for GTA activity, and the active fractions were pooled for dialysis, by using Slide-a-Lyzer dialysis cassettes, against 2 × 4 liter of 50 mM MOPS, 1 mM DTT, 100 mM NaCl, 5 mM $MnCl_2$, pH 7.0, to remove the UDP. The samples that had a low activity were concentrated in a centrifuge using a Centriplus 30 filtration unit. After the dialysis, the samples were assayed for activity and protein concentration by using the modified assay and mass quantitation methods described below. The homogeneity of protein was analyzed by SDS-PAGE as described (18).

3.2.4 Standard assays for GTA activity

The enzyme was diluted in a 2 × A/B buffer (2 mg/ml BSA, 40 mM $MnCl_2$, 100 mM MOPS, pH 7.0), and 5 μl of this dilution was added to 5 μl of GTA assay mix (140 μM Fucα1–2Galβ-O-$(CH_2)_7CH_3$ acceptor, 140 μM UDP-GalNAc donor, and 0.029 μCi UDP-[6-^{3}H]GalNAc donor (approximately 64,000 dpm)). The incubation time was 30 min, at 37 °C, and the amounts of acceptor and donor consumed were less then 10% of

the MAX. At the end of the incubation time, the reaction was quenched with 0.5 ml cold deionized water, and the reaction product was transferred to a C_{18} cartridge (Sep-Pak) on a vacuum manifold. The cartridge was previously washed with 50 ml MeOH then 50 ml deionized water. After sample loading the cartridge was washed with 50 ml deionized water, and the radio-labeled product was eluted with 3.5 ml MeOH into a scintillation vial. Liquid scintillation cocktail (10 ml) was added to each vial for counting (5 min) in a Beckman liquid scintillation counter (LS 1801).

3.2.5 Standard assays for GTB activity

The enzyme was diluted in a 2 × A/B buffer (2 mg/ml BSA, 40 mM $MnCl_2$, 100 mM MOPS, pH 7.0) and 5 μl of this dilution was added to 5 μl of GTB assay mix (990 μM Fucα1–2Galβ-O-$(CH_2)_7CH_3$ acceptor, 532 μM UDP-Gal donor, 0.02 μCi UDP-[6-^{3}H]GalNAc donor (approximately 44,000 dpm)). The incubation time was 30 min, at 37 °C, and the amounts of acceptor and donor consumed were less then 10% of the MAX. At the end of the incubation time, the reaction was quenched with 0.5 ml cold deionized water and was treated as described for the standard GTA assay.

3.2.6 Saturating assays

The modified assay mix contained sufficient acceptor and donor to fully saturate the enzymes. These were used to assess the final activity of the respective enzymes.

3.2.6.1 Saturating assays for GTA activity

The enzyme was diluted in a 2 × A/B buffer (2 mg/ml BSA, 40 mM $MnCl_2$, 100 mM MOPS, pH 7.0), and 5 μl of this dilution was added to 5 μl of GTA assay mix (1200 μM Fucα1–2Galβ-O-$(CH_2)_7CH_3$ acceptor, 2000 μM UDP-GalNAc donor, and 0.086 μCi UDP-[6-^{3}H]GalNAc donor (approximately 190,000 dpm)). The incubation time was 30 min, at 37 °C, and the amounts of acceptor and donor consumed were less then 10% of the MAX. At the end of the incubation time, the reaction was quenched with 0.5 ml cold deionized water and was treated as described for the standard GTA assay.

3.2.6.2 Saturating assays for GTB activity

The enzyme was diluted in a 2 × A/B buffer (2 mg/ml BSA, 40 mM $MnCl_2$, 100 mM MOPS, pH 7.0), and 5 μl of this dilution was added to 5 μl of GTB assay mix (1200 μM Fucα1–2Galβ-O-$(CH_2)_7CH_3$ acceptor, 2000 μM UDP-Gal donor, and 0.02 μCi UDP-[6-^{3}H]GalNAc donor (approximately 50,000 dpm)). The incubation time was 30 min, at 37 °C, and the amounts of acceptor and donor consumed were less then 10% of the MAX. At the end of the incubation time, the reaction was quenched with 0.5 ml cold deionized water and was treated as described for the standard GTA assay.

3.2.7 Protein determination

The Bio-Rad protein assay procedure based on Bradford's method (19) was used to estimate protein concentrations. Bovine gamma globulin was used as a protein standard at 0.05, 0.1, 0.2, 0.4, 0.6 mg/ml.

3.2.8 Kinetic characterization

Assays were carried out at 37 °C in a total volume of 15 μl containing the substrates and enzyme in a 50 mM MOPS buffer, pH 7.0, with 20 mM $MnCl_2$ and 1 mg/ml BSA. Seven different concentrations of the donor or acceptor were used, and the amount of substrate consumed was less than 15% to ensure linear initial reaction rates. Data were analyzed by fitting to the Michaelis-Menten equation (Eq. 1).

$$\upsilon = \frac{V_{\max}[S]}{K_m + [S]} \qquad \text{(Eq. 1)}$$

The parameter [S] in this equation denotes acceptor or donor substrates. The kinetic parameters V_{max} and K_m were derived from the best fit of the Michaelis-Menten equation using nonlinear regression with the GraphPad Prism 3.0 program.

3.2.9 Enzymatic synthesis of A and B structures by L266G

3.2.9.1 A-product

The preparative scale enzymatic synthesis for the A-product was carried out in a reaction mixture containing 1.1 mg (2.55 μmol) of precursor H oligosaccharide Fucα1–2Galβ-O-$(CH_2)_7CH_3$ acceptor, 2.3 mg (3.53 μmol) of UDP-GalNAc, 0.5 mg/ml BSA, 1.6 μl (1.6 units) of alkaline phosphatase, 2.6 units of concentrated L266G *cis*-AB enzyme, 50 mM MOPS buffer, pH 7.0, in a total volume of 500 μl The reaction mixture was incubated for 180 h at 37 °C. The progress of the reaction was monitored by thin layer chromatography (TLC) with CH_2Cl_2-MeOH-H_2O (65:35:6, v/v) as a solvent system. When the reaction was complete by TLC, product was isolated on two Sep-pak C_{18} reverse phase cartridges as described previously (20). The product was characterized by ^{1}H NMR spectroscopy on an INOVA *i*600 spectrometer at 600 MHz and by mass spectrometry on a PerSeptive Biosystems Mariner Biospectrometry Workstation.

3.2.9.2 B-product

The enzymatic synthesis of the B-product was carried out in a reaction mixture containing 0.7 mg (1.28 μmol) of precursor H oligosaccharide Fucα1–2Galβ-O-$(CH_2)_7CH_3$ acceptor, 2.9 mg (5.04 μmol) of UDP-Gal, 0.5 mg/ml BSA, 1.0 μl (1 unit) of alkaline phosphatase, and 0.7 units of concentrated L266G *cis*-AB enzyme, 50 mM MOPS buffer, pH 7.0, in a total volume of 320 μl The reaction was incubated for 180 h at 37 °C and was treated as described for synthesis of A-product in 3.2.9.1.

3.2.10 Crystallography

The L266G mutant was crystallized using conditions similar to the native GTA and GTB enzyme (4). Data were collected on a Rigaku R-AXIS4++ detector at a distance of 100 mm and exposure times between 4.0–5.0 min for 0.5 degree oscillations. X-rays were produced on a MicroMax 30 W generation system coupled to Osmic "Blue" confocal x-ray mirrors. The crystals were frozen at the cold head and maintained under cryogenic conditions at a temperature of -160 °C. Data sets for L266G unliganded and in

complexes with UDP-GalNAc or UDP-Gal and acceptor were solved using native GTA or GTB (Protein Data Bank accession code 1LZ0, 1LZ7, respectively) as a starting model with refinement using the program CNS (21) and modeling with the software SETOR (22).

3.3 Results

3.3.1 Site-directed mutagenesis

The L266G mutant was prepared by two rounds of PCR using the AAAB clone as a template. The first PCR was performed to make short and long fragments and the subsequent PCR was necessary to make the full gene. The gel of each product of the first PCR revealed that the size of both short and long fragments were correct (approximately 270 and 620 bp, respectively) (Fig. 3–4). In the second PCR, the short and long products were combined giving an approximately 900 bp gene product (Fig. 3–5). The purified PCR product and pCWΔlac were digested and then combined for the ligation reaction. Plasmids purified from individual *E. coli* transformants were digested with *Eco*RI and

FIG. 3–4. **Agarose gel of the first PCR fragments for the L266G mutant.** The left gel shows the short fragment, the right gel the long fragment.

FIG. 3–5. **Agarose gel of the second PCR reaction for the L266G mutant.**

*Hin*dIII (Fig. 3–6). The lower band of Fig. 3–6 confirmed the correct size of digested L266G product from the purified plasmid (approximately 900 bp).

3.3.2 Expression and purification

The soluble form of the L266G *cis*-AB mutant was expressed in *E. coli* and purified on a SP-Sepharose FF ion-exchange column and a UDP-Hex affinity column. The purified enzyme, crude extracts, and all of the flow-through from the columns were monitored by SDS-PAGE (Fig. 3–7). The lanes represent the crude lysate after French Press disruption (Lys), the supernatant from ultracentrifugation (S/N), the eluant from the SP-Sepharose FF column from loading the enzyme (SP-LFT), the eluant from the SP-Sepharose FF when washing the column with equilibration buffer (SP-WFT), the eluant from the SP-Sepharose FF column when washed with equilibration buffer containing 0.5 M NaCl (SP-eluant), the eluant from the UDP-Hex column when loading the enzyme (UDP-LFT), the eluant from the UDP-Hex column when washing the column with equilibration buffer (UDP-WFT) and the enzyme after the desorption from the UDP-Hex

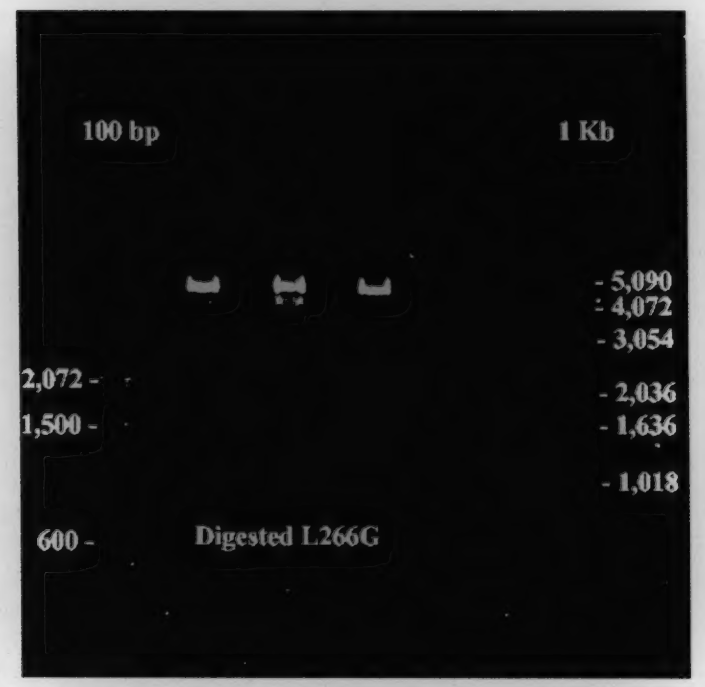

FIG. 3–6. **Agarose gel of digested L266G mini plasmid.** Upper and lower bands digested with *Eco*RI and *Hin*dIII represent pCWΔlac plasmid and the L266G mutant, respectively.

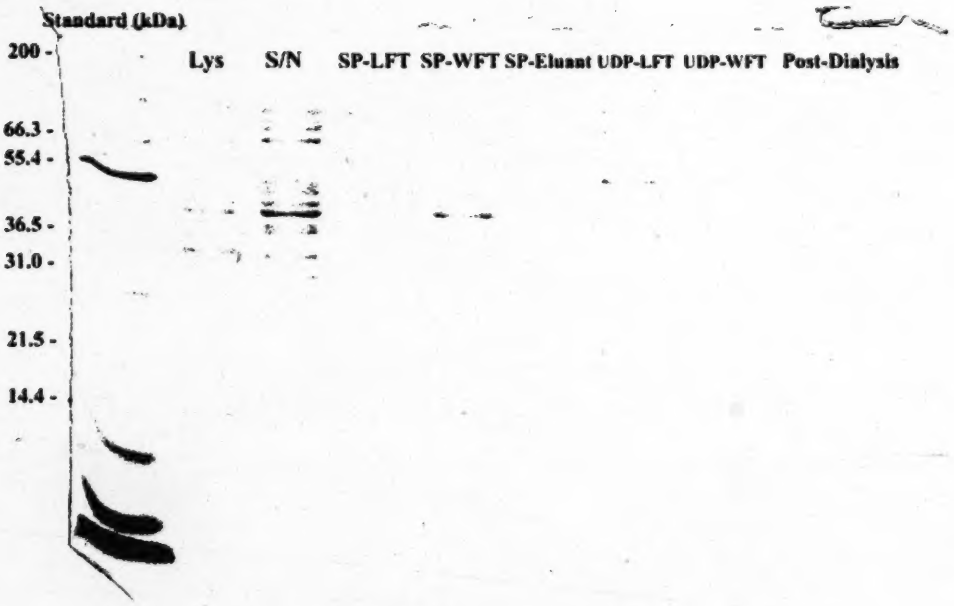

FIG. 3–7. **SDS-PAGE result for L266G mutant.**

column and dialysis to remove UDP (Post-dialysis). Molecular mass standards are also shown on the gels. The specific activity of purified *cis*-AB was 0.9 Unit/mg using the saturating GTA assay and 0.3 Unit/mg using the saturating GTB assay where a unit is the amount of enzyme that converts 1 μmol of substrate to product per minute at 37 °C. The yield of the final purified protein was 30 mg/liter of *E. coli* cell culture.

3.3.3 Kinetic characterization

AAAB was the template to make the *cis*-AB mutant (L266G). The hybrid AAAB enzyme showed a 4-fold higher k_{cat} than L266G (AAGlyB) using UDP-GalNAc and 10-fold higher k_{cat} using UDP-Gal as donor (Table 3–1). The 3-fold higher k_{cat} of this hybrid enzyme for UDP-Gal compared to the k_{cat} value for UDP-GalNAc revealed that AAAB is a B-like transferase. The K_m for acceptor and donor were all lower than those for AAGlyB.

The AABB mutant also exhibits the dual functionality but, showed less transferase activity than AAAB. Its k_{cat} values with UDP-GalNAc were 0.6 s^{-1} and 2.2 s^{-1} with UDP-Gal donor. Because AABB had a very low K_m for acceptor and donor, it

TABLE 3–1.

Kinetic constants for wild-type and cis-AB glycosyltransferases

Enzyme	Donor	K_A[a]	K_B[b]	k_{cat}	k_{cat}/K_A	k_{cat}/K_B
		μM	μM	s^{-1}	$mM^{-1}s^{-1}$	$mM^{-1}s^{-1}$
GTA	UDP-GalNAc	9.9	8.7	18	1770	2010
GTB	UDP-Gal	88	27	5.1	58	190
AAAB	UDP-GalNAc	16	13	2.4	149	185
	UDP-Gal	45	12	7.6	169	633
AABB	UDP-GalNAc	13	37	0.6	46	16
	UDP-Gal	1.6	1.8	2.2	1375	1222
AAGlyB[c]	UDP-GalNAc	82	248	0.66	8.1	2.7
	UDP-Gal	42	95	0.78	19	8.2

[a] K_A is the Michaelis-Menten constant for the acceptor (Fucα1–2Galβ-O-$(CH_2)_7CH_3$).

[b] K_B is the Michaelis-Menten constant for the donor (UDP-GalNAc or UDP-Gal).

[c] AAGlyB mutation was based on the AAAB clone as a template.

showed higher k_{cat}/K_m values especially in the reaction with the B donor.

The L266G mutant showed a 26-fold lower k_{cat} compared to that of the wild-type GTA using UDP-GalNAc as a donor and a 6.5-fold lower k_{cat} value compared to that of the wild-type GTB using UDP-Gal as a donor. This mutant showed higher K_m values compared to those of the wild-type GTA(B) enzymes. The K_A was 8 times greater than wild-type GTA but only two times greater than wild-type GTB. K_B values showed the same pattern, they were much larger than wild-type GTA (28 times) and only 3.5 times larger than those of wild-type GTB. The specificity constant (k_{cat}/K_m), which combines the effects of both rate and binding, was decreased 218-fold for the acceptor, and 755-fold for the donor using the A donor (UDP-GalNAc), and showed 3-fold increase for the acceptor and 23-fold decrease for the donor using B donor (UDP-Gal) (Table 3–1). However, the k_{cat} for both donors were almost identical, indicating that this mutant (AAGlyB) has equal dual functionality.

3.3.4 Enzymatic synthesis

The kinetic data suggest that the L266G mutant enzyme can be used with UDP-

Gal or UDP-GalNAc to synthesize A and B antigens (Fig. 3–1). In preparative synthesis reaction progress was monitored by TLC during incubation. The product GalNAcα1–3[Fucα1–2]Galβ-O$(CH_2)_7CH_3$ and Galα1–3[Fucα1–2]Galβ-O$(CH_2)_7CH_3$ have an R_f of 0.54 whereas the starting material Fucα1–2Galβ-O$(CH_2)_7CH_3$ has an R_f of 0.87. The respective products were obtained in good yield (Table 3–2) and their structures were confirmed by their ^{1}H NMR (Fig. 3–8, 3–9) (Table 3–3) and mass spectral data (Table 3–4).

3.3.5 Crystal structure

The crystal structure of the L266G provides a structural basis for its reduced enzymatic activity. Table 3–5 shows the statistical results of the L266G mutant grown unliganded and in complex with acceptor and UDP-GalNAc or UDP-Gal. All structures showed excellent electron density over the polypeptide chain, with the exception of the disordered loop (aa 176–195) and the final 10 amino acid residues at the C-terminal, which were also absent in the native glycosyltransferase structures. The active site region of L266G mutant enzyme showed excellent electron density (Fig. 3–10).

TABLE 3–2.

Enzymatic synthesis reactions of L266G

Donor	Starting	Enzyme	Volume	Reaction time	Yield	Yield
	mg	*unit*	*μl*	*h*	*mg*	*%*
A[a]	1.1	2.6[b]	500	180	1.5	99
B[c]	0.7	0.7[d]	320	180	0.6	60

[a] A denotes UDP-GalNAc

[b] Unit of enzyme calculated with respect to the standard assay for GTA with UDP-GalNAc

[c] B denotes UDP-Gal

[d] Unit of enzyme calculated with respect to the standard assay for GTB with UDP-Gal

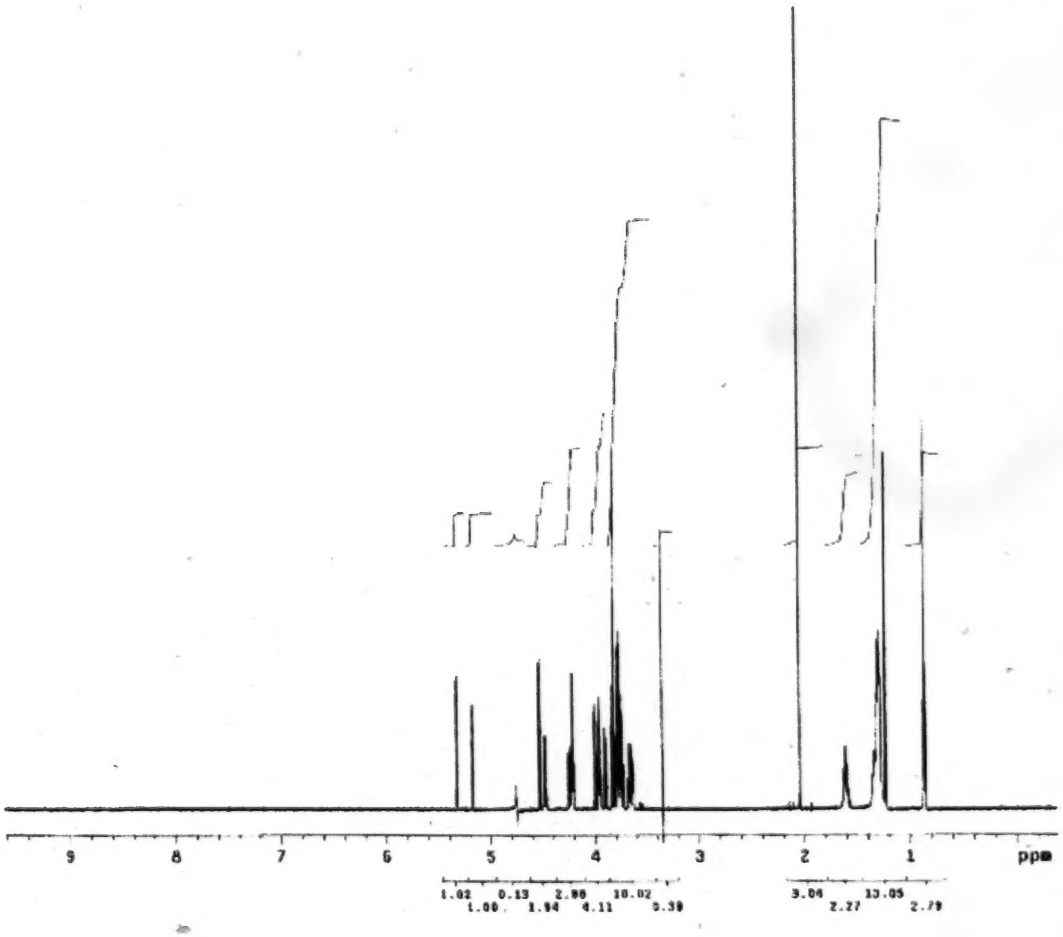

FIG. 3–8. ^{1}H NMR spectra (in D_2O) of the A-product (GalNAcα1–3[Fucα1–2]Galβ-O-$(CH_2)_7CH_3$).

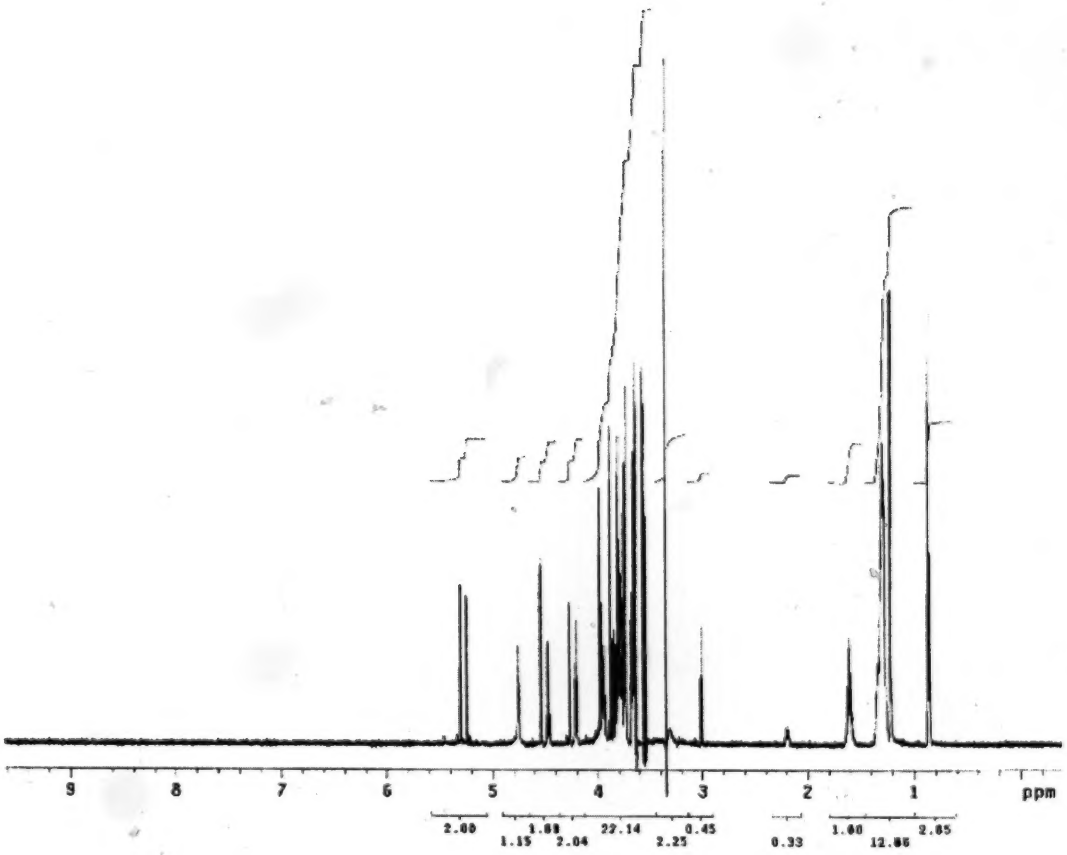

FIG. 3–9. ^{1}H NMR spectra (in D_2O) of the B-product (Galα1–3[Fucα1–2]Galβ-O$(CH_2)_7CH_3$).

TABLE 3–3.

Selected ^{1}H NMR data for trisaccharide products of L266G

	A-product		B-product	
	Donor			
	GalNAc		Gal	
	Chemical shift *ppm*	Coupling constant *Hz*	Chemical shift *ppm*	Coupling constant *Hz*
Gal/GalNAc α-H-1	5.24	3.0	5.17	3.6
NHAc	2.04	N/A		
GalNAc α-H-2	4.24	3.6 & 11.4		
Gal β-H-1	4.54	7.8	4.52	7.8
Gal β-H-4	4.20	6.6	4.21	4.8
Fuc α-H-1	5.30	3.6	5.33	4.2
Fuc α-H-5	4.47	6.4	4.47	6.7
Fuc α-H-6	1.22	6.6	1.23	6.6
Octyl CH_3	0.86	6.6	0.86	6.6

TABLE 3–4.

ESMS data for the trisaccharide products of L266G

Reaction for	Calculated for	Observed Mass	Calculated Mass
		g/mol	*g/mol*
A - product	$C_{28}H_{51}NO_{15}Na$	*664.3156*	*664.3151*
B - product	$C_{26}H_{48}O_{15}Na$	*623.2885*	*623.2885*

TABLE 3–5.

Data collection and refinement statistics for L266G mutant grown in the unliganded and presence of donors and acceptor

	L266G	L266G + HA + UDP-GalNAc	L266G + HA[a] + UDP-Gal
Resolution (Å)	20–1.8	42.06–1.44	19.81–1.49
Space group	$C222_1$	$C222_1$	$C222_1$
a (Å)	52.5	52.5	52.7
b (Å)	149.0	149.0	149.0
c (Å)	79.3	79.3	79.3
R_{merge} (%)[b,c]	4.5 (31.8)	5.8 (30.2)	3.6 (24.8)
Completeness (%)[b]	91.4 (50.6)	86.8 (21.0)	92.9 (54.4)
Unique reflections	46,979	49,014	47,619
Refinement			
Resolution	20–1.8	20–1.8	20–1.8
R_{work} (%)[d]	20.1	20.3	20.1
R_{free} (%)[e]	22.5	23.0	22.3
r.m.s.[f] bond (Å)	0.005	0.005	0.005
r.m.s. angle (°)	1.29	1.29	1.29

[a] HA, H-acceptor

[b] Values in parentheses represent high resolution shell.

[c] $R_{merge} = \Sigma|I_{obs} - I_{ave}|/\Sigma I_{ave}$

[d] $R_{work} = \Sigma||F_o| - |F_c||/\Sigma|F_o|$

[e] 10% of reflections were omitted in R_{free} calculations.

[f] r.m.s., root mean square.

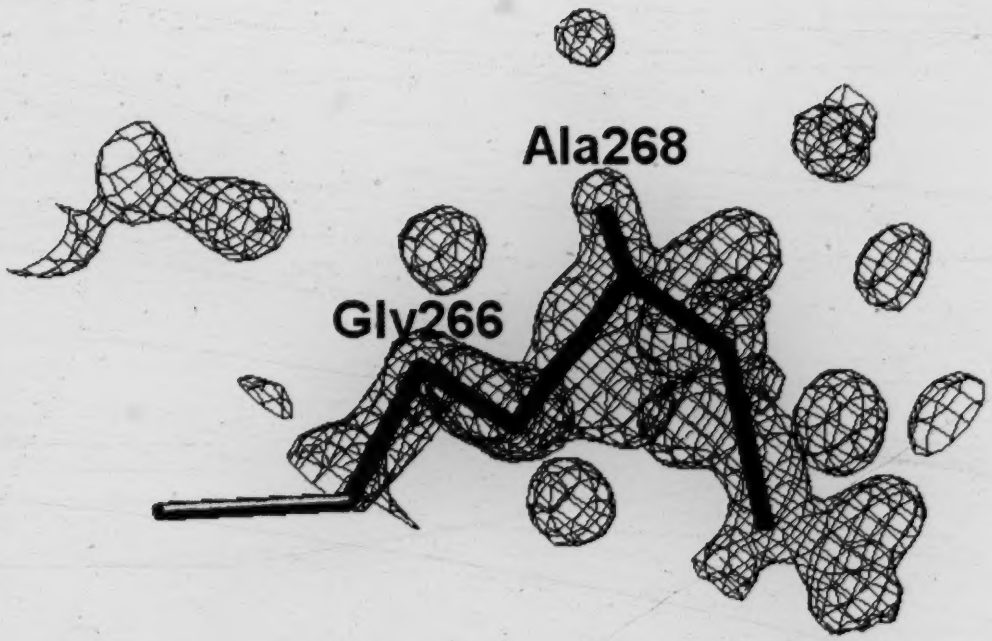

FIG. 3–10. **The electron density map of active cleft of the L266G mutant.** Substitution of L266G enlarges the active site cleft and allows binding of both the UDP-GalNAc and UDP-Gal donors. *Red* represents Ala-268. Gly-266 is not coded by color due to the absence of a side chain. Limited main chains (264–270) are shown. The electron density map was represented by SETOR (22).

3.4 Discussion

In this chapter the production of a mutant AB glycosyltransferase by site-directed mutagenesis and expression of *cis*-AB glycosyltransferase (AAGlyB) was carried out. This *cis*-AB enzyme showed a high K_m constant compared to that of the wild-type GTA and GTB. Its K_A for the B donor (UDP-Gal) was 3-fold higher than that of the wild type GTB, indicating weaker binding. The AAAB enzyme which is used a template to make the *cis*-AB mutant in this study showed higher k_{cat} than that of AAGlyB. This indicates that the amino acid at position 266 has a significant role in the enzymatic reaction. A study of the AABB mutant revealed that the k_{cat} of this enzyme was lower than that of the AAAB mutant. The only difference between these hybrid enzymes is an amino acid alteration at position 266, either a Leu in AAAB or Met in AABB. Because only Leu/Met-266 is positioned to contact the acetamido/hydroxyl group of the A/B donor it distinguishes between UDP-GalNAc and UDP-Gal donor. This is a complementary interaction where the larger acetamido group of the A donor is accommodated by the smaller Leu-266 in the GTA, whereas the smaller hydroxyl group of the B donor is

accommodated by the larger Met-266 in GTB (4). Therefore, the change of Leu-266 alters the substrate specificity.

In the *cis*-AB glycosyltransferase (AAGlyB) the mutation at the residue 266 to Gly creates larger space in the binding pocket of the enzyme. This enlarged binding site allows a high degree of motion to the UDP-Gal substrate and enough space to bind with UDP-GalNAc; that is, this mutant can react with both the UDP-GalNAc and UDP-Gal. However, this enzyme (AAGlyB) showed a higher K_A with the A donor (UDP-GalNAc) than that of AAAB or AABB, indicating a defect in binding ability due to the mutation. The glycine residue at 266 creates a larger space for a donor; as a result, the sugar binding moieties are altered. This alteration makes a larger active site cleft that lacks of complementariness with the A donor (UDP-GalNAc).

The k_{cat} constant of the AAGlyB is 27 and 7 times lower than that of wild-type GTA and GTB, respectively and also lower than that of AAAB and AABB. The high value of K_m probably contributes to this low k_{cat} because of the incomplete binding of the donor. Glycine is a small residue that induces an improper fit with a donor. As a result, the transferase activity of the *cis*-AB enzyme (AAGlyB) is decreased. However, the two k_{cat} constants for the A donor and B donor are almost identical, indicating that the

AAGlyB has a dual functionality. The specificity constant (k_{cat}/K_m) also showed lower values than those of wild-type GTA, GTB, AAAB, and AABB implying that transferase efficiency is decreased due to the mutation at 266. Both the high K_m and low the k_{cat} contributed to the low enzyme efficiency.

The crystal structure confirms that the substitution of AAGlyB creates the enlarged active site cleft. Modeling approaches explain the dual functionality and represent the position of the substrates in the active site cleft of the AAGlyB (Fig. 3–11). Because of the lack of the side chain at 266, the *cis*-AB cannot discriminate between donors efficiently. This result is reasonable because Leu/Met-266 dominates A/B donor specificity (16, 23) and because these residues at 266 is positioned to contact the characteristic acetamido/hydroxyl groups (4). As a result of the L266G mutation, the *cis*-AB glycosyltransferase does not distinguish well between UDP-GalNAc and UDP-Gal. In addition, the enlarged active site cleft allows accepting both donors. However, the modified active site reduces the specificity for donors because the active site becomes more flat, therefore, interferes with the complementary interaction with donor substrates (Fig. 3–12). This is the reason for the lower k_{cat} values of the *cis*-AB enzyme.

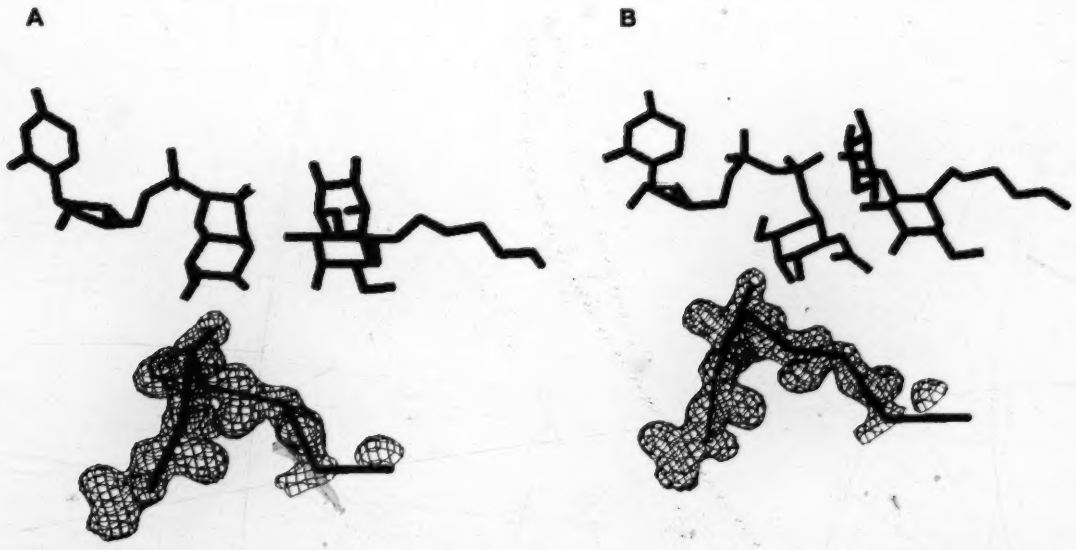

FIG. 3–11. **The electron density map of the active cleft of the L266G enzyme (AAGlyB) with donor substrates.** The enlarged active site cleft of the L266G mutant enzyme allows accepting both UDP-GalNAc and UDP-Gal donors. *Red* represents Ala-268. Gly-266 is not represented by color due to the absence of the side chain and limited main chains (264–270) are shown. UDP-GalNAc and UDP-Gal are represented by *yellow*. *Blue* represents the H-acceptor. Fig. 3–11–*A* is the modeled structure with H acceptor and A donor (UDP-GalNAc) and Fig. 3–11–*B* is the modeled structure with the same acceptor and B donor (UDP-Gal). The GalNAc and Gal moieties were modeled by SETOR (22).

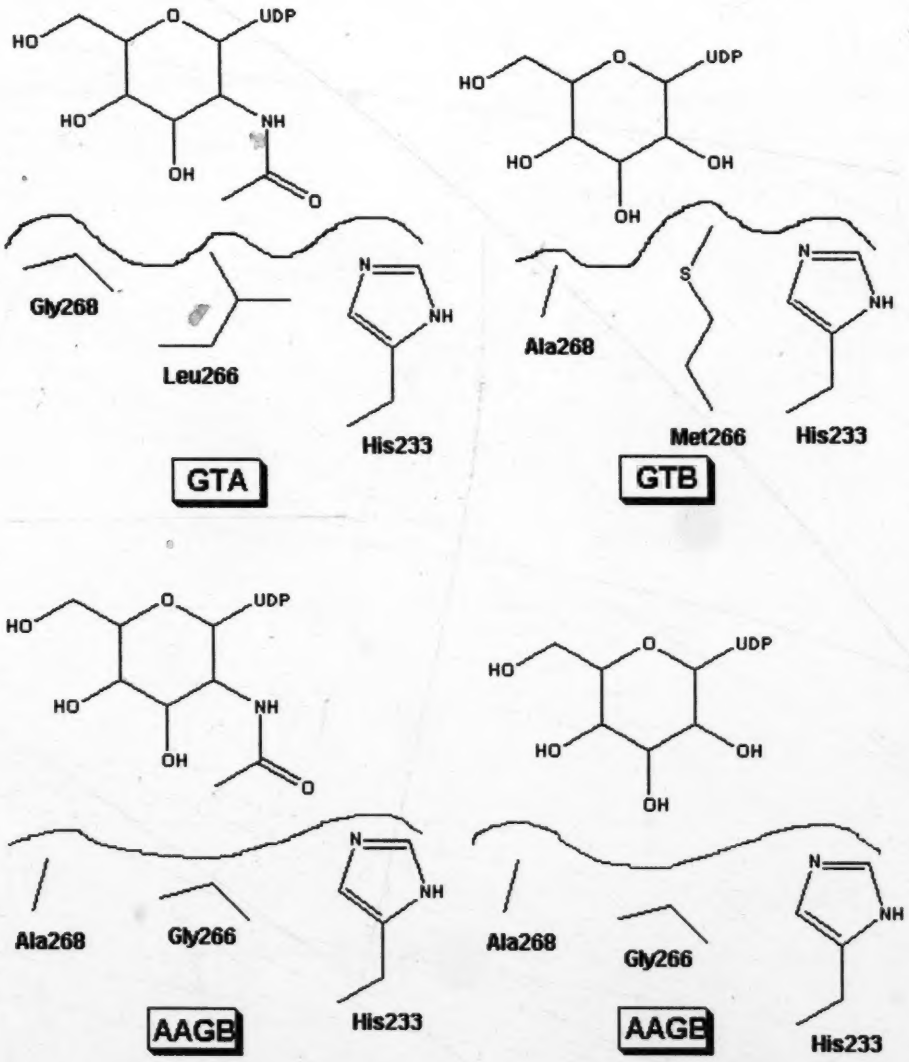

FIG. 3–12. **Schematic representation of the active site of the *cis*-AB enzyme (AAGlyB) with donor substrates.** The enlarged active site floor of the AAGlyB allows accepting both donors (UDP-GalNAc and UDP-Gal), however it induces a lack of complementary interactions with donor substrates. Fig. 3–12–*A* represents a schematic view of the active sites of GTA and GTB with their characteristic donors; UDP-GalNAc and UCP-Gal, respectively. Fig. 3–12–*B* represents the active site of the *cis*-AB glycosyltransferase (AAGlyB). The *green* lines denote the flattened active site of the *cis*-AB with UDP-GalNAc and UDP-Gal, respectively.

^{1}H NMR data were analyzed to confirm characteristic protons of A and B products synthesized by AAGlyB; donor α-H-1, GalNAcα-H-2, Galβ-H-1, Galβ-H-4, Fucα-H-1, Fucα-H-5, Fucα-H-6 and $(CH_2)_7CH_3$ (Table 3–3). The A product showed the terminal GalNAc in an $\alpha(1\rightarrow3)$ linkage to Gal, and the B product showed the terminal Gal in an $\alpha(1\rightarrow3)$ linkage to Gal as reported previously (24). The mass spectral data confirmed that a GalNAc or Gal residue was added to the Fucα1–2Galβ-O-$(CH_2)_7CH_3$ acceptor (Table 3–4).

Acknowledgements

Assistance with DNA sequencing was provided by the Molecular Biology Service Unit, Department of Biological Sciences, University of Alberta. Assistance with NMR and mass spectra was provided by the NMR and Mass Spectroscopy Laboratory in the Department of Chemistry, University of Alberta. We thank F. W. Dahlquist for his gift of a pCW vector, W. Wakarchuk for a pCWΔlac vector, and O. Hindsgaul for the Fucα1–2Galβ-O-$(CH_2)_7CH_3$ acceptor. This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC), the Alberta Ingenuity Centre for Carbohydrate Science (AICCS) and the Canadian Institute of Health Research (CIHR).

References

1. Watkins, W. M. (1980) *Adv. Hum. Genet.* **10,** 379–385

2. Palcic, M. M., Seto, N. O., and Hindsgaul, O. (2001) *Transfus. Med.* **11,** 315–323

3. Yamamoto, F., Clausen, H., White, T., Marken, J., and Hakomori, S. (1990) *Nature* **345,** 229–233.

4. Patenaude, S. I., Seto, N. O. L., Borisova, S. N., Szpacenko, A., Marcus, S. L., Palcic, M. M., and Evans, S. V. (2002) *Nat. Struct. Biol.* **9,** 685–690

5. Yamaguchi, H., Okubo, Y., and Hazama, F. (1965) *Proc. Jpn. Acad.* **41,** 316–320

6. Yamamoto, F., McNeil, P. D., Yamamoto, M., Hakomori, S., Bromilow, I. M., and Duguid, J. K. (1993) *Vox Sang.* **64,** 120–123

7. Fukumori, Y., Ohnoki, S., Yoshimura, K., Nagao, N., Shibata, H., Tomica, T., Okubo, Y., and Yamagochi, H. (1996) *Transfus. Med.* **6,** 337–344

8. Ogasawara, K. Yabe, R., Uchikawa, M., Saitou, N., Nannai, M., Nakata, K., Takenaka, M., Fujisawa, K., Ishikawa, Y., Juji, T., and Tokunaga, K. (1996) *Blood* **88,** 2732–2737

9. Mifsud, N. A., Watt, J. M. Condon, J. A., Haddad, A. P., and Sparrow, R. L.

(2000) *Transfusion* **40,** 1276–1277

10. Yamamoto, F., McNeil, P. D., Yamamoto, M., Hakomori, S., and Harris, T. (1993) *Vox Sang.* 64, 171–174

11. Yu, L. C., Lee, H. L., Chan, Y. S., and Lin, M. (1999) *Biochem. Biophys. Res. Commun.* **262,** 487–493

12. Yamamoto, M., Lin, X., Kominato, Y., Hata, Y., Noda, R., Saitou, N., and Yamamoto, F., (2001) *J. Biol. Chem.* **276,** 13701–13708

13. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (eds) (1997) *Current Protocols in Molecular Biology*, Vol. 1, pp.8.5.7–10, John Wiley and Sons, Inc., New York

14. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, New York

15. Seto, N. O. L., Palcic, M. M., Compston, C. A., Szpacenko, A., and Palcic, M. M. (2000) *Carbohydr. Res.* **324,** 161–169

16. Seto, N. O. L., Compston, C. A., Evans, S. V., Bundle, D. R., Narang, S. A., and Palcic, M. M. (1999) *Eur. J. Biochem.* **259,** 770–775

17. Gegner, J. A., and Dahlquist, F. W. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 750–

754

18. Seto, N. O. L., Palcic, M. M., Hindsgaul, O., Bundle, D. R., and Narang, S. A. (1995) *Eur. J. Biochem.* **234,** 323–328

19. Bradford, M. M. (1976) *Anal. Biochem.* **72,** 248–254

20. Palcic, M. M. (1994) *Methods Enzymol.* **230,** 300–316

21. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sec. D.* **54,** Pt. 5, 905–921

22. Evans, S. V. (1993) *J. Mol. Graph.* **11,** 134–138

23. Seto, N. O. L., Palcic, M. M., Compston, C. A., Li, H., Bundle, D. R., and Narang, S. A. (1997) *J. Biol. Chem.* **272,** 14133–14138

24. Lemieux, R. U., Bock, K., Delbaere, L. T. J., Koto, S., and Rao, V. S. (1980) *Can. J. Chem.* **58,** 631–653

Chapter 4. Future work

For the past several years numerous ABO alleles have been genetically characterized. Comparison of these ABO alleles allows the identification of the amino acid sequence differences that alter the specificities and activities of the translated blood group glycosyltransferases. Despite this genetic information on ABO alleles, most enzymes encoded by these alleles have not yet been purified, kinetically characterized or structurally analyzed.

The glycosyltransferases encoded by the weak A or B alleles including A^3, A^x and B^3 can be constructed by site-directed mutagenesis using GTA or GTB as a template (Fig. 4–1, 4–2). The enzymes translated by other alleles that contain extended reading frames such as A^2 and A^{el} may be constructed and purified using similar experimental procedures as those described in Chapters 2 and 3.

Other glycosyltransferases encoded by the blood group alleles which show *cis*-AB behaviors such as AAAB, BBAB, BABB, ABAB, and BAAB (Fig. 3–3) can be constructed by the identical biochemical techniques used in this thesis.

Part of the nucleotide sequence of the A^x and $B^{(A)}$ alleles has not been determined.

After the genetic basis of these alleles is determined, new types of glycosyltransferases that are encoded by these alleles can be constructed, purified and characterized.

The O^3 glycosyltransferase (Fig. 4–1, 4–2) can be purified by similar methods described in previous chapters. The O^3 enzyme may not be purified easily due to its non-functionality. Therefore, alternative experimental procedures may be required. For example, protein concentration can be measured consecutively during the purification to find fractions of interest. Crystallographic analysis of the O^3 enzyme with the corresponding donor and acceptor can help to find the electron density of GalNAc and Gal moieties that were not visible in previous trials.

In the future all of the glycosyltransferases that are encoded by the potential ABO alleles described above may be constructed, purified and kinetically characterized. Crystallographic analysis can then be used to reveal the detailed structural features of the translated enzymes.

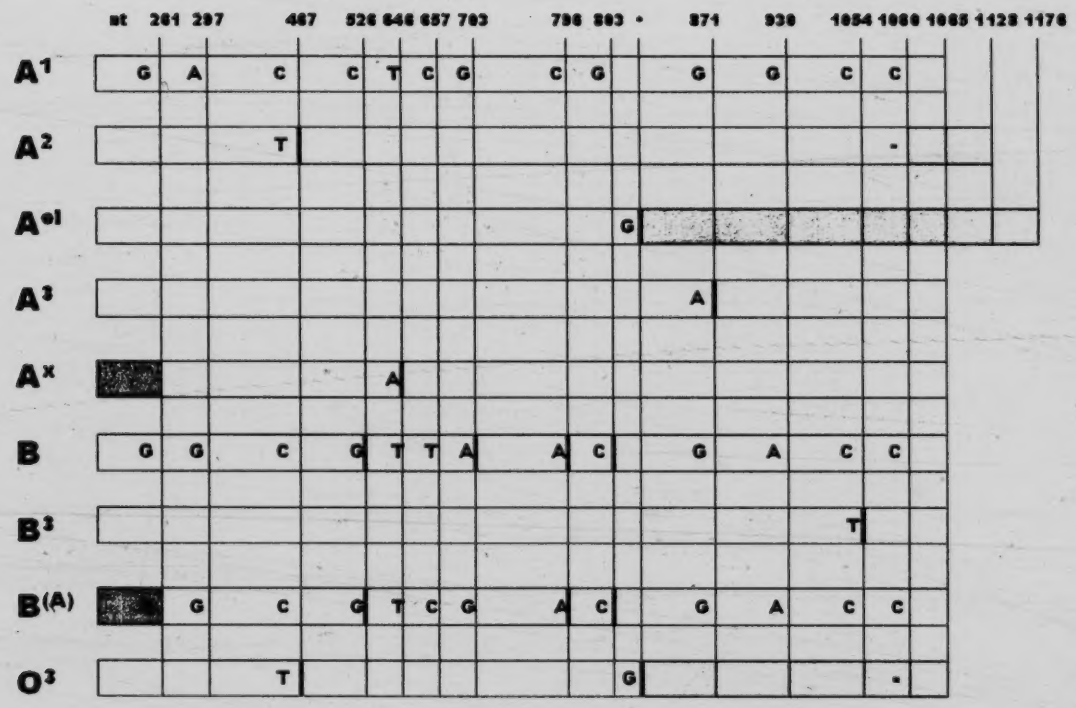
nt 261 297 467 526 646 657 703 796 803 871 939 1054 1060 1065 1123 1176
A1 G A C C T C G C G G G C C
A2 T -
Ael G
A3 A
Ax A
B G G C G T T A A C G A C C
B3 T
B(A) G C G T C G A C G A C C
O3 T G -

FIG. 4–1 **The nucleotide sequence of the ABO blood group alleles which have not yet been characterized.** Mutations causing amino acid changes are represented by red vertical bars. The negative (-) sign denotes the deletion of the nucleotide and the asterisk (*) indicates the guanosine insertion at the nt. 798–804 of the A^{el} and O^3 allele. The orange areas indicate the entirely different deduced amino acid sequence in alleles due to reading frame shift. The green indicates the nucleotide sequence has not been determined.

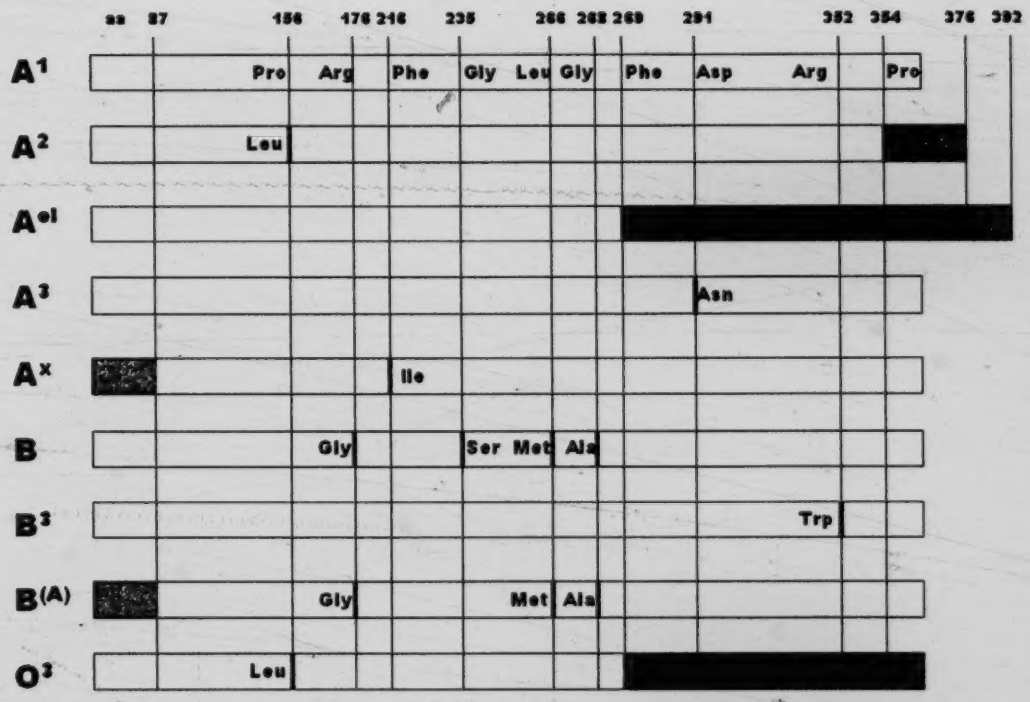

FIG. 4–2. **The amino acid sequence of the ABO blood group alleles which have not yet been characterized.** Mutations are represented by red vertical bars. The orange areas indicate the entirely different amino acid sequence due to reading frame shift. The green indicates the amino acid sequence has not yet been determined.